OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.4_p5_4578

Run on: April 15, 2003, 00:17:03 ; Search time 2426.96 seconds
                 (without alignments)
                 10930.628 Million cell updates/sec

Title: US-09-001-737-7_COPY_15_1652
Perfect score: 1638
Sequence: 1 ATGGCAAAAGAAATCAAATT..........TGGGTGGGATGGGCGGATAA 1638

Scoring table: IDENTITY.NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 listing first 45 summaries

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estmu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: em_estl:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est1:*
                 13: gb_est3:*
                 14: gb_est4:*
                 15: gb_est5:*
                 16: em_estfun:*
                 17: gb_gss:*
                 18: em_gss_hum:*
                 19: em_gss_inv:*
                 20: em_gss_pln:*
                 21: em_gss_vrt:*
                 22: em_gss_fun:*
                 23: em_gss_mam:*
                 24: em_gss_mus:*
                 25: em_gss_other:*
                 26: em_gss_pro:*
                 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	471.8	28.8	2265	11	AY109623	AY109623 zea mays
2	431.4	26.3	2290	11	AY104969	AY104969 zea mays
3	430.6	26.3	2338	11	AY103778	AY103778 zea mays
c 4	414.2	25.3	1327	17	BH770613	BH770613 LLMGtag37
5	383.6	23.4	2143	11	AY108560	AY108560 zea mays
6	302.8	18.5	1376	11	AY106465	AY106465 zea mays

RESULT 1
AY109623

LOCUS	AY109623		2265 bp	mRNA		linear		HTC	25-MAY-2002

DEFINITION Zea mays CL2221_1 mRNA sequence.
ACCESSION   AY109623
VERSION   AY109623.1 GI:21213414
KEYWORDS   HTC.
SOURCE   Zea mays.
  ORGANISM Zea mays.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 2265)
  AUTHORS   Rainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE   Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2 (bases 1 to 2265)
  AUTHORS   Coe,E.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
FEATURES   Location/Qualifiers
     source   1..2265
                     /organism="Zea mays"
                     /db_xref="MaizeDB:630067"
                     /db_xref="taxon:4577"
                     /clone="CL2221_1"

ALIGNMENTS

7	290	17.7	1387	11	AY108518	AY108518 zea mays
8	263.2	16.1	887	17	CNS074DM	CNS074DM
9	262.8	16.0	1090	14	AJ428672	AJ428672 clone BA0
10	261.8	16.0	878	12	BM799922	BM799922 AGENCOURT
11	257.2	15.7	1035	9	BF275584	BF275584 GA__Ed002
12	253.4	15.5	988	9	AL518632	AL518632 AL518632
13	253	15.4	977	12	AL532233	AL532233 AL532233
14	251.6	15.4	955	9	BG321293	BG321293 Zm04_04g0
15	249.6	15.2	1017	9	AL515262	AL515262 AL515262
16	249	15.2	677	13	AL515579	AL515579 AL515579
17	248.6	15.2	810	12	BI926411	BI926411 EST546300
18	248	15.1	959	14	BF627437	BF627437 HVSMEb000
19	247	15.1	746	14	AL557150	AL557150 QGG35F17.
20	246.4	15.1	980	9	BU003474	BU003474 QGG35F17.
21	245	15.0	761	14	AL557181	AL557181 AL557181
22	243	14.8	645	10	BG967621	BG967621 OHB30G09
23	242	14.8	633	14	AW224051	AW224051 EST300862
24	239.4	14.6	716	14	BU014013	BU014013 QG6D07.Y
25	237.6	14.5	637	10	D46006	D46006 RICS10372A
26	237	14.5	1119	13	AW223768	AW223768 EST300579
27	236.2	14.4	668	10	BM458421	BM458421 AGENCOURT
28	233	14.2	771	12	AW398404	AW398404 EST298251
29	231.2	14.1	677	10	BF617440	BF617440 HVSMEcO01
30	230.2	14.1	698	12	AV938920	AV938920 AV938920
31	229.2	14.0	1051	13	BM478208	BM478208 AGENCOURT
32	228.8	14.0	744	14	BU000648	BU000648 QGG25I09.
33	227.4	13.9	680	12	AW330455	AW330455 PFOVAFCB3
34	227.2	13.9	632	12	BG128308	BG128308 EST473954
35	226.6	13.8	650	9	AJ432272	AJ432272 AJ432272
36	226	13.8	790	10	BG599254	BG599254 EST504154
37	226	13.8	1351	10	BE422296	BE422296 HWMO22cF.
38	225.6	13.8	895	14	BQ216655	BQ216655 AGENCOURT
39	225.4	13.8	946	14	BQ882339	BQ882339 AGENCOURT
40	225.2	13.7	709	9	AI486676	AI486676 EST244998
41	225.2	13.7	1016	13	BM450130	BM450130 AGENCOURT
42	224.8	13.7	739	14	BQ915279	BQ915279 QHB14A23.
43	223	13.6	744	13	BI176450	BI176450 EST517395
44	222.8	13.6	654	13	BI176166	BI176166 EST517249
45	222.2	13.6	660	12	BG354884	BG354884 MBTsMLAl3

/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

BASE COUNT   615 a   430 c   599 g   538 t   83 others
ORIGIN

Query Match   28.8%;   Score 471.8;   DB 11;   Length 2265;
Best Local Similarity   56.4%;   Pred. No. 1.7e-120;   Matches 911;   Conservative   0;   Mismatches   694;   Indels   9;   Gaps   3;

```
Qy      4 GCAAAAGAAATCAAATTTTCAGCAGATGGCGGTGCGTGCCACGATGGTCGCGGAGTTGATATG   63
Db    293 GCCAAGGAGACATCAAGTTTGGTTGGTTTGAGCGCCGCTCTCGATGNNNNNNNNNNNNNNNNNNN  352

Qy     64 TTAGCAGATACCGTCAAAGTAACCGTTGGTCCTAAAGGCGGCAATGTTGTTGTTCTTGAAAA  123
Db    353 NNNNCAGATGCTGTCAAAAGTGACAAATGGGACCTAAGGGGGCCAAATGTGTTATTGAGCAA  412

Qy    124 GCTGTTTGGTTCTCCCTTAATTACTAATGACACGGGGTAACCATTGCTAAAGAGATCGAATTA  183
Db    413 AGCTTGGGTGCACCGAAAGTCAACAAAGGATGGTGACTGTACCAAAGAACATTGAATTT  472

Qy    184 GAAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTCTCAAGGTGGCCTTCTAAAACCAAT  243
Db    473 AAGGATAGAGTAAAGAAATGTTGGGTCGCAAGCCTTGTTGAAAACAGGTTGCTAATGCAACTAAT  532

Qy    244 GATATTGCTGGTGATGGGAGGACGACTACTGCCAACAGTTTTGACACAGCCCATTGTTCAGA  303
Db    533 GACACTGCTGCGGGATGGTACCACATGTGCCACTGTTTTGACAAAAGCAATATATTACTGAG  592

Qy    304 GGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGTATCCGTCGAGGCATTGAAACA  363
Db    593 GGGTGCAAAATCTGTTGCGCGGGCTGAATGAATCCAATGGATTTAAGGCGTGGAATCTCAAATG  652

Qy    364 GCAACAGCAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCTGTATCTGCAAGGAA  423
Db    653 GCTGTTGACGCTGTTGTGTGACCAATCTCGAAAGGCATGGCCAGAATGATCACTTCAGAA  712

Qy    424 GCTATTGCTCAGTCGCCTGCCGAGTATCATCACGCTTCTGAAA---AAGTTGGAGAGTATATC  480
Db    713 GAAATTGCACAGGTGGGTACAATATCAGCAAATGGGAAAGGGAAATTGGTCGAGCTTATT  772

Qy    481 TCAGAAGCTATGGAGGGTGTGGGCCACGAGATGGTGTGATTACCATCGAAGAATCTCGAGGT  540
Db    773 GCCAAGGCTATGGAGGAAGGTTGGCAAAGAGAGGTGTCATCACTATTGCGGACGGTAACACC  832

Qy    541 ATGGAAAACAGAACTCGAAGTGGTTGAAGGGCATCGCAAATTGACCGTGGTTGACCTGTCTCACA  600
Db    833 CTTTATATAATGAGCTTGAAGTTGTGGAGGTATGAAAACTAGACAGAGGTTACAYCTCTCCG  892

Qy    601 TACATGGTCACAGACAACAATGAAAAAAATGGTTCCAGACCTTGAAAACCCATTATNCTTAATC  660
Db    893 TACTTCATTACCAACTCAAAAGCCCAGAAAATGTGAATTGGAAGACCCATTGATCTTAATT  952

Qy    661 ACGGATAAAAAAGTGTCAAAACATCCAAGAGACAATTTTGCCACTACTTGAGGAAGTTCTTAAA  720
Db    953 CATGACAAGAGAGGTTACGAATAATGCATGCTGTGGTTAAGGTTTTAGAAATGGCCTCTGAAG  1012

Qy    721 ACCAACGTTCCATTACTACTCATTATTGCAGATGATGTGGGATGGTGAAGCACTTCCAACCCTT  780
Db   1013 AAACAAAAAGCCTCTACTGATTGGTTGCAGAGATGTGAAGCAATTGGGCACTTTG  1072

Qy    781 GTCTTGAACAGAATTCGTGTGACTTTCAATGTGGTTGGCTGTTCAAAGCGCCAGGATTTGGT  840
Db   1073 ATTATTACAACACCTTCGTCGCAGGCATCAAGGTCTGTCTGTCTGTCAAAGCTTCCTGGTTTTGGG  1132

Qy    841 GATCGTCGTAAAGCTATGCTTGAAGACAATTGCTATCTTGACAGGTTGTTACAGTGATTACA  900
```

Db 1133 GAAAACAGGAAGGCAAACTTACAGGACCTTGCAAATCCTTACTGGAGAGAGAGGTAATAACT 1192
Qy 901 GAGGATCTAGGACTTGAATTAAAAGATGCTACAAATGACAGCCCTTGGGACAGGCGTGCTAAG 960
Db 1193 GAAGAACTAGGAATGAACCTTCAGAATTTGAGCCTCACATGCGGTCACATGCCNNNNNG 1252
Qy 961 ATTACAGTTGATAAAGATAGCACACAGTAATTGTTGAAGGTTCAGGAGTTCAGAGGCTATT 1020
Db 1253 GTGACTGTCTCTAAGGATGACACTGTTATTCTTGATGAGGCCGGACAAGCAAGAAGTCCATT 1312

Qy 1021 GCTAACCGTATTGCCACTGCATTAAATCGGCAATTAGAAACAACAACTTCTGACTTTGACCGT 1080
Db 1313 GAAGAGAGGGCCACAGACAGAGTTAGAATCAGCAATTGACAACTTCAGATTATNGATTAAG 1372

Qy 1081 GAAAAACTACACAAACAGTTTTGGGCAAATTAGCTCGGTGTGGCTGTTATCAAAGTAGGA 1140
Db 1373 GAAAAGCTCCAGGAGCCGGTTGGCAAAGCTCTCCTGGAGGTGTTGCCTTCTTCTGAAGATTGGA 1432

Qy 1141 GCTCCAACAGAGACAGCAGCTTTAAAAGAAATGAAAACTTCCCATTGAGGAGATGCTCAAATGCT 1200
Db 1433 GGAGCCAGCGAAGCAGAAGTTGGTGAGGAAGGATTAGAGTGACAAGCACTTATTAGGGTT 1492

Qy 1201 ACACGTGCAGCCGTTGAAGAAGGTTATCGTTCCTGGTGGTGGAAACAGCACTTATTACGGTT 1260
Db 1493 ACTAAAGCTGCTCGTTGAAAGAAGGTTATTGTTACCAGGTGGTGTTGTGCTTCTCTCATTGGCA 1552

Qy 1261 ATTGAAAAAGTAGGTCAGCGCTCTTGAGCCTTG---AGGGCGGATGAATGGCTACTGGAGGTAACATT 1317
Db 1553 TCGAAGGAGGCTTGATAAAATTGCACAACAGGCAAACTTCGATCAGAAGAGATGGGTTGTTGCAAAAATC 1612

Qy 1318 GTGGCTTCGTGCCTCTACAGAAGAGCCTGTACGTCAAAATGCCTTGGTTGGTTAAATGCTGGGTACCGAAGGC 1377
Db 1613 ATTCAGAAATGCTTTGAAGAGAAAGCTTTTGCGACCAAGAAAATACTGACCTGGGTTTATGATGCTGCT 1672

Qy 1378 TCCGTAGTTATTGACACAAG---TTGCAAAAAACAGCCCTGCAGGAACAGGAGATTTAATCCTTCGA 1434
Db 1673 GCAGTAGTTGTTGTAGGAAAAGCTTTTGGACCAAGAAAATACTGACCTGGGTTTATGATCGTGCT 1732

Qy 1435 ACAGGTGAGTGGGGTTGATATGATTAAAACAGGAATCATTGACCCTTGTCAAAGTAACACACGA 1494
Db 1733 AAAGGTGAATAATGTTGACATGGTGAGACACGGTATCATTCAGATGAGGTTGATCAGA 1792

Qy 1495 TCAGCGCCTTCAAAATGCAGCTTCTTGTTGTGAGCTAGTCTTAATTTGACAAAACAGAGAGCAGTTGTT 1554
Db 1793 ACTGCTTTGGTGGATCGTAGTGTGTTCCTGATGACAACCACCACGGGAATCCAATAATT 1852

Qy 1555 GCTAATAAAACCTGAACCAGCTACGGCCAGCCAGCAGTGCCAGCAGGTGATGGAT 1608
Db 1853 GTAGAGAATTCCCAAGGAACAGAGGCACCGGGCTCCCAGCAATGGGTGGCGGCCATGGGT 1906

RESULT 2
AY104969
LOCUS   AY104969   2290 bp   mRNA   linear   HTC 25-MAY-2002
DEFINITION   Zea mays PCO070942 mRNA sequence.
ACCESSION   AY104969
VERSION   AY104969.1   GI:21208047
KEYWORDS   HTC.
SOURCE   Zea mays.
ORGANISM   Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
           clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 2290)
AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
           Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE   Maize Mapping Project/DuPont Consensus Sequences for Design of
           Overgo Probes
JOURNAL   Unpublished (2002)
REFERENCE   2 (bases 1 to 2290)
AUTHORS   Coe,E.C.
TITLE   Direct Submission
JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of

```
FEATURES        Location/Qualifiers
source          1..2290
                /organism="Zea mays"
                /db_xref="MaizeDB:633867"
                /db_xref="taxon:4577"
                /clone="PC0070942"
                /clone_lib="Maize Mapping Project/DuPont Cornsensus
                Library"
                /note="this sequence is part of a project of EST
                assemblies resulting from the application of public
                contigs to seed DuPont contigs; this resource was
                assembled by DuPont as part of a collaboration for the
                overgo addressing of BACs in conjunction with the Maize
                Mapping Project."
BASE COUNT      684 a     457 c     583 g     565 t     1 others
ORIGIN
```

```
Query Match           26.3%;   Score 431.4; DB 11;  Length 2290;
Best Local Similarity 57.9%;   Pred. No. 3.6e-109;
Matches 883; Conservative     0; Mismatches 522; Indels  19; Gaps  6;
```

```
Qy    52 GGAGTTGATATGTTAGCAGAGATACCGTCAAAGTAACCCTTGGTCCTAAAGGGCGCAAATGTT 111
         || ||  |||| ||  ||| | | ||| ||| |||||   ||  || |||||||| |||||||
Db   426 GGGGTCAACAAGCTTGCAGATCTCGTCGGGGTTACTCTTGGACCAAAAGGGGCGAAAATGTA 485

Qy   112 GTTCTTGAAAAAGCTTTTGGTTCTCCCTTAATTACTTAAATGACGGGGTAACCATTGCTAAA 171
         |||| || ||||     |||| ||||| || |||| |    || ||| || | |||| |||
Db   486 GTCCTCGAAAGCAAGCTATGCGTTCACCTAAAATTGTTCAATGATGGGGTGTCACTGTTGGCAAAG 545

Qy   172 GAGATCGAATTAGAAGATCATTTTGAAAAACATGGGAGCAAAATTGGTTGTGTCTGAAGTGGCT 231
         || || |  ||  | ||  ||| |||||| | ||  |||||||| | |||| ||||| || ||
Db   546 GAGGTTGAACTTGAGGAGCCCTTCTTTGAAAATATTGGTGGCTAAAATTGGTCCCCCAAGCTCT 605

Qy   232 TCTAAAACCAATGATAATTGCTTGGTTGGTGGTGACGGACTACTCCAACACTGTTTTGACACAGGCC 291
         ||| |    || |  || |  |||  ||||  ||||   ||| || ||| | ||   | ||| ||
Db   606 GCTAAAACAAATGACTTGGCGTGTCAATGGGACAACCCACTTCGGTTGTTCCTTGCTCAAGGG 665

Qy   292 ATTGTTCATGAAGGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGTATCCGGTCGA 351
         |||   ||||| |||| ||    || || | ||| |||||||||| | ||  ||| |||
Db   666 ATGATTACTCAGGGTGTAAAGGTTGTGACAGGCAGTGTGCTAACCCTGTAATCCAGGATCACCCGT 725

Qy   352 GGCATTGAAACAGCAACAGCCAACAGCTGTGTTGAAGCCATTGCTCAACCTGTA 411
         ||| |||| ||| || || | ||  |||||  |||| || |  |||| || ||
Db   726 GGTATTGAGAAAAACAGCCAAAGCACTAGTCAGTGAACTCCAAAAGAGTGTCCAAGGAGGT 785

Qy   412 TCTGGGCAAGGAAGGTATTGCTCAGGTCCCTGCCAGTATCATCACGCCTCT---GAAAAAGTT 468
         |||  |||| || ||  ||  | |  |||   ||  |  || || | ||||     |||||| |
Db   786 TGAGGACAGTGAGC---TTGCTGATGTTCCTCAGTTGGTTAGTGCCGGGAAATAACTATGAAAATT 843

Qy   469 GGAGAGTATATCTCAGAAGCTATTGGAAGCTGCTGTGGGCACGATGGTGGTTGATTACCCATCGAA 528
         ||   | |   | || ||| |  || || | | |||  | | | ||  |||||| |  | || |||
Db   844 GGGAATATGATAGCTGAGGCTATGAAAAGGTTGGGCGGCCAGGGAGGTGTCAACACTTGAA 903

Qy   529 GAATCTCGAGGTATGGAAACAGACTTGAAGTGGTTGAAGGCATGCAATTTGACCGTGGT 588
         |||   || | |  ||| |||| |  |  |  ||  |  || |  ||  |  || |||
Db   904 GAAGGAAAGAGTGCTGAAAATAACCTCTATTCTGTTGANGGAATGCAGTTTGACCGTGGC 963

Qy   589 TACCTGTCTCAAATACATGGTCACAGAACAATAAAAAATGGTTGCACAGACCTTGAAAACCCA 648
         |||| |  |||  || | | ||| | || |    ||||||  ||  ||| | ||| |  || |
Db   964 TATATTTTCTCCCTACTTTGTGACTGACAGTGAAAAAATGATCAGTGCGAGTATGAGAATTGC 1023

Qy   649 TTTATCTTAATCACGGATAAAAAAGTGTCAAACATCAAGACACATTTTGCCACTACTTGAG 708
         |  |  | |  |  | | || |||||  |  || |  || | | ||| |||| | || ||
Db   709 GAAGTTCTTAAAACCACCGTTCCATTACTTACTTCATTATTGCAGATGAGTTGAAGGTGAAGGCA 768

Qy  1024 AAGCTTCTCCTTGTTGGACAAGAAAAATTAACACATTTAACTTATCACTACCTGGAG 1083
         ||  | | |||  |||  |  |   | | |  |  |  |  |   || | |  ||||
Db  1084 GATGCTATTAGAAGTGGATAACCCAATTCTAATAGAGTGGCAGAGGACATTGAGCAGGAAGCT 1143

Qy   769 CTTCCAACCCTTCTTCTTGACAAGATTCGTCGTACTTTCAAATGTGGTTGGCTGTCAAAGCG 828
         | |||   |  |   |  | | ||  ||   | |   | |||| | |  | || | ||||
Db  1144 CTTGCAACCCTTGTGGTTAATAGGCTTAGAGGTTGCATTGAAGATTGCTGCTTATTAAGGCC 1203

Qy   829 CCAGGATTTGGTGATCGTCGTGTAAAGCTATGCCTTGAAGACACATTGCCTATCTTGACAGGTGGT 888
         || || ||  ||  | | ||  || ||  ||| |||| ||||| ||| ||  |   |  ||| |
Db  1204 CCTGGTTTTGGAGAGCGGCGAAAAGTCAAAATATCTTGATGCAATTCGTACTCTCACTGGAGGC 1263

Qy   889 ACAGTGATTACAGAGGATCTAGGACTTGAATTAAAAGATCTACAAATGACAGCCCTTGA 948
         ||| |  |||| || ||| | |||  ||||  |||| || || |  ||| || | ||
Db  1264 ACTGTCATCACAGAGAAGAAGAGTTGGATTATCCTTAGACAAGGCACAACAATGGGGTCCTAGGA 1323

Qy   949 CAGGCTGCTAAGATTACAGTTGATAAAGATAGGCACACAGTAATTGTTGAAGGTTCAGGAGT 1008
         ||| ||  ||  |||  | | ||||| |  |  |||| ||  | | |||| || |||||| |
Db  1324 ACTGCTCCTAAGGTTGTTAGTTACCAAGAGATTCAAACAATATAGTGGAGATTGGGCACTACA 1383

Qy  1009 TCAGAAGCTATTGCTTAACCGTATTGCCACTGATTAAATCCCAATTAGAAACACAACTTCT 1068
         |  ||| ||  | |  | ||  |  |  || |   || ||||| | |  || ||  |
Db  1384 CAGGAGGAAGTGAACAAAAGGGTGACACAAAATCAAGACAAACCAGAATTGAGGGCTACTGAACAA 1443

Qy  1069 GACTTTGACCGTGAAAAACTACAAGAAACGTTTGGGCGAAATTACCTGTGTGGTGTAGCTGTT 1128
         || |  ||  || ||| | ||  ||| | |  ||| || || ||| | | ||| |  |  |
Db  1444 GAATACGAGAGAGAAAAACTAAACGAGAGTAGCTAAACTCTCTGGTGGTGGTTGCTGTC 1503

Qy  1129 ATCAAAGTACGGAGGCTCCACAGACAGACACGCTTTAAAAGAAAATGAAAACTTCCCATTGAGGAT 1188
         ||  | | ||  |||| ||| || |  |  | ||| |  |||  | ||||  ||||| |  |||||
Db  1504 ATTCAGGTAGGAGCACAGCAGAAACTGAGCTCAAGGAAAAAGAAAATTGAGGGTTGAGGAT 1563

Qy  1189 GCTCTAAATGCTACACGTGCGACCGGTTGAGAGAAGGTATCCGTTGCTGTGGTGGGAACAGCA 1248
         ||  |  ||| |  |  | ||  | ||| ||| | | ||  | ||| | | ||  | |||
Db  1564 GCCCTGAATGCTACAAAGGCAGCACGTCAGGCAGAAGGTATCGGTTGTAGGTGGTGGCTGTGGTACC 1623

Qy  1249 CTTATTACGGTTGATTATTGAAAAAGTAGCACGCTCTTGAG------CTTGAGGCGCGATGAGTGCT 1302
         || ||   ||  ||      ||| || | || | ||||||        ||||| || ||| | |
Db  1624 CTTTTGAGGCTTGCATCAAAAAGTTGCAATTATAGAGACAGCCCTTGGCGAACGAAGGACAG 1683

Qy  1303 AC---TGGACGTAACATTGTCGTCCTCTAGAAGCAGCCCTGTACGTCAAAATTGCTTTA 1359
         | |  ||  || ||| || | |  | |  | ||| || ||| | | ||  |||  ||||
Db  1684 AAGGTTGGGCGTGAAATTGTAGGAGAGGTCCTTGACGTATCCACTTAAATTGATTGCAAAA 1743

Qy  1360 AATGCTGGGTACGAAGGGTCCGTAGTTATTGACAAG----TTGAAAACAGCCCTGCGAGG 1415
         |   |  |  ||  | | ||  | | | || || ||    ||| |||| || ||  ||||
Db  1744 AATGCAGGCGTTAATGGCAGTGTCGTTTACTGAGAAGGTTCTTGCCAATGAGAGAACTTCAGG 1803

Qy  1416 AACAGGATTTAAATGCTCGCAACAGTGTGAGTGGGTTTGATTAGATTAAAACAGGAATCATTGA 1475
         |  ||| ||  |  | | |  || ||  |  |  |  ||||| |  |  |||| ||||||||||
Db  1804 TATGGG-TACAATGCGCGCTACAGGGGAGTATGAGGACTTGATGGTCGTATCATTGA 1862

Qy  1476 CCCTGTCAAAGTAACACGCATCAGCCGCTTCAAAATGCAGCGTTCTGTGACCTAGTGCTTATTTT 1535
         ||| | |||| |  |  | || ||||  || |||   || ||  | |||  | | | |  ||||
Db  1863 TCCCACCAAGGTTGTGAGGTGCTGCTGGAGCAGTCGCTGGGCGAAAACCTTCAT 1922

Qy  1536 GACAACAGAAGCAGTTGTTGCTAA 1559
         || ||   |||| || || | |||
Db  1923 CACCTCTGATGCCGTTGTTGTTGA 1946
```

```
RESULT 3
AY103778
LOCUS           AY103778               2338 bp    mRNA         linear   HTC 25-MAY-2002
DEFINITION      Zea mays PC0063180 mRNA sequence.
ACCESSION       AY103778
VERSION         AY103778.1   GI:21206856
KEYWORDS        HTC.
SOURCE          Zea mays
  ORGANISM      Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE       1  (bases 1 to 2338)
  AUTHORS       Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
                Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE         Maize Mapping Project/DuPont Consensus Sequences for Design of
                Overgo Probes
  JOURNAL       Unpublished (2002)
REFERENCE       2  (bases 1 to 2338)
```

```
AUTHORS   Coe,E.C.
TITLE     Direct Submission
JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
          Missouri, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1..2338
                     /organism="Zea mays"
                     /db_xref="MaizeDB:633377"
                     /clone="PC063180"
                     /clone_lib="Maize Mapping Project/DuPont Corngenous
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT    668 a    483 c    643 g    544 t
ORIGIN

Query Match       26.3%;  Score 430.6;  DB 11;  Length 2338;
Best Local Similarity   56.2%;
Matches   875;  Conservative    0;  Pred. No. 6.2e-109;  Mismatches  669;  Indels  13;  Gaps  3;
```

```
Qy    52 GGAGTGATATGTTAGCAGATACCGTCAAAGTAACGCTTGGTGCCTAAAGGGCGCAATGTT 111
Db   452 GGAGTCAATAAGCTTGCAGACCTAGTTGGAGTTACACTCGGACCAAAAGGGAAGAAGTGTC 511

Qy   112 GTTCTTGAAAAAGCTTTTGGTTCTCCCTTATTACTAATGACGGGGTACCAATTGCTAAA 171
Db   512 GTTTTGGGAGGAGCAAGTACGGCTCTCCCAAAGATTGTTAACGAATGGGGTTACACGTTGCAAGA 571

Qy   172 GAGATCGAATTAGAGAGATCATTTTGAAAACATGGGAGCAAAATTGGTGTGTCTGAAGTGGCT 231
Db   572 GAGGTTGAGCTGGAGAGACCCTGTCGGAAAAACATTGGAGCTAAATTGGTCAGGCAAGCTGCA 631

Qy   232 TCTAAAACCAATGATAATTGCTGGTGGTGAGTGGGACGACTACTGCAACAGTTTTGACACAAGCC 291
Db   632 GCTAAGACCAACGAATCTTGCTGCAGATAGGACAACCACCTCTGTGGTCCTTGCTCAGGGG 691

Qy   292 AATGTTCATGAGAGGACTAAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATCCGTCGA 351
Db   692 CTGATTGCTGAGGGTGTTAAGGTTGTGGCAGCCTGGTGCTAATCCCTGTTCAGATTACTCGT 751

Qy   352 GGCATTGAAACACACAACAGCTGTTGAAGGCCTTGAAAGCCATTGCTCAACCTGTA 411
Db   752 GGTATCGAGAAACACGAAAAGCACTAGTTGAAGAACTACGGAAGTTGTCTAAGGAGGTT 811

Qy   412 TCTGGCAAGGAAGCTATTGCTCAGGTCGCTGCAGTATCATCACGCTCTGAAAAGTTGGA 471
Db   812 GAAAGATAGTGACCTTGCAGATGTTGTCGTCGTTAGTCGTGCTGGCACAACTATGAAATTGCT 871

Qy   472 GAGTATATCTCAGAAGCTATGGAGGCGTCGGGCAACGATGGTGTGATTACCATCGGAAGAA 531
Db   872 AACATGATTAGCAGAGGCCATGAGCAAGGTCGGTCGGGTTGGGTTACCCTCGAAGAG 931

Qy   532 TCTCGAGGTATGGAAAACAGAACTTGAAGTGGTTGAAGGCCATGCAAATTGACCGGGTTAC 591
Db   932 GGGAGGAGTTCTGAGGAACTTCTCCTATGTTGTTGGAGGGAAATGCAAATTTGACCGTGGTTAC 991

Qy   592 CTGCTCCAATACATAGGTCACACACATGAAAAAATGTTGCAGACCTTGAAAAACCCATTT 651
Db   992 ATCTCCGTACTTTGTTACAGACAAGTGCAGAAAATGTCTGCCGAGTATATGAGAACTCCAAG 1051

Qy   652 ATCTTAAATCACGGATAAAAAAAGGTGTCAAACATCCAAGACATTTTGCCCAC--TACTTGCAGGA 710
Db  1052 CTGCTTTGGTTGACAAAAGGATCACCAAGGGTCAGCTTACGTTTTGGAGA 1111

Qy   711 AGTTCTTAAAACCAACGTGCCATTACTTCATTTATTTGCAGAATGATGTTGACAGGACACT 770
Db  1112 AGCCATCAGGAGGTTGCATACCCAATCCTGATAATTGCTTGAGGAGATATTGAGCAGGAGGCTCT 1171

Qy   771 TCCAACCCTTGTCTCTTGAACAAGAATTCGTGGGTTACTTTTCAATGTGGTTGCTGTAAAACGCGCC 830
Db  1172 TGCTACCCTTGTCTGCTCCTCAACAAGCTCAGAAGGATCACTGAAGATCGCTGCCCATCAAAACCCC 1231

Qy   831 AGGATTTGGTGATCGTCGTAAAGCTATGCTTGAAGACATTGCTATCTTGACAGGTGGTAC 890
Db  1232 TGGTTTTGGAGACGCCAAGACTCAGTACTTGGATGGATAATTGCCATCCTTACTGAGCGAAC 1291

Qy   891 AGTGATTACAGAGGAGGATCTAGGGACTTGAATTAAAAAAAGATTGCTAGACAGCACGCCTTGGAGA 950
Db  1292 TGTAATCAGAGATTGAGAAGAGTTGGGCGTCATTGCAGCACAAGGGTGACCAAGGGCTGAC 1351

Qy   951 GGCTGCTAAGATTACACTTGATTAAAAGATAGCCACAGTTAATTGTTGAAGGTTCAGGAAGTTC 1010
Db  1352 AGCTGCAAAAGGTTGTCCTTACGGAAAGAGTCGGACAACAATTGTTGGTGATTGGCAGCAGCCCAA 1411

Qy  1011 AGAAGCTATTGCTCAACCGTATTGCACTGATTAAAATCGCAATTAAAAATTCATTGAGGAACAACAACTTCTGA 1070
Db  1412 GGAAGAAGTTACTAAGAGGGGTTGCACAGATTAAAAAATTCATTGAGGCAGCAGACAGA 1471

Qy  1071 CTTTGACCGTGAAAAACTACAGAAGAACGTTTGGCGGAAATTAGCGTGGTGGGTGTAGCTGTTAT 1130
Db  1472 ATATGAAAAGGAAAAACGCTCAATGAGGAGGATAGCAAAAGCTTCGTGCGTGGTGTTGCTGTTCAT 1531

Qy  1131 CAAAGTAGGAGGAGTCCCAACAGACAGACAGCTTTAAAAGAGAATGAAACTTCCCAGTTGAGGAGTCC 1190
Db  1532 TCAGGTGGGAGGCACAAAACAGAAACTGAACTAAAAGAGAAGAAGTTAAGAGTTGAGACGATCC 1591

Qy  1191 TCTAAAATGCTACACGTGCCACGCCGTTGAAGAAGGTATCCTTGCTGGTGTCGTGGAAACAGCACT 1250
Db  1592 CCTAAATGCAACCAAGGCTGCTGCTTGTTGAGGAAGGTTATGTCTTGGTGGTGGACATCCCT 1651

Qy  1251 TATTACGGTTGATTGAAAAAGTAGCAGTCTCTTGAGCTT-----GAGGCGGATGATGC 1301
Db  1652 TTTGCGGCTTGCAGCTGAAAAGTCGCAATGCCATCACAGGGATACTCGGAACGATGAGCAGAGA 1711

Qy  1302 TACTGGACGTAACATTGTTGCTTCCGTGCTCTAGAAGAGGCCTGTACGTCAAATTGCTTTAAA 1361
Db  1712 GTTTGGACGTTGCGAAATAGTTAAGGAGGGCCCTGAGCTACCCACTTAAATTGAATTGCTAAAAA 1771

Qy  1362 TGCTGGGTACGAAGGCTCCGTAGTTATTGACAAGTTGAAAAAACAGCCCTGCGAGGAA---C 1418
Db  1772 TGCCGGTGTCAATGGCAGCCGTCGTCACTGAGACAGGTCCTTCTTAACGATAACTTCAAGTA 1831

Qy  1419 AGGATTTAATGCTGCAACAAGGTGAGTGGGGTTGATATGATCATTAAAAACAGGAAATCATTGACCC 1478
Db  1832 TGGTTACAATGCTGCTACTGGACAGTACGAGGACTTGATTGGCTGGTGTATCATTGATATCC 1891

Qy  1479 TGTCAAAGTAACACAGTCACGGCTTCAAAATGCAGCGTTCTTGCTAGCTACTCTTATTTTGAC 1538
Db  1892 TACTAAGGTGGTGAGATGCTGTTTGGAGCAATGCCGCCAAGAGCTTCCTGAC 1951

Qy  1539 AACAGAAGCAGTTGTTGCTAAATAAAACCTGAACCAGTTACCGCGCACCAGAAATGCC 1595
Db  1952 ATCTGATGTTGTGTGGTCGTCGGATATTAAGGAGCAGAGCCGGAGAGTCTGCGCCCCTTGCTAACCC 2008
```

```
RESULT 4

BH770613/c

LOCUS       BH770613                1327 bp    DNA     linear   GSS 01-MAY-2002
DEFINITION  LLMGtag9372 MG1363 Random Sequence Tag Library Lactococcus lactis
            subsp. cremoris genomic, DNA sequence.
ACCESSION   BH770613
VERSION     BH770613.1  GI:20373570
KEYWORDS    GSS.
SOURCE      Lactococcus lactis subsp. cremoris
  ORGANISM  Lactococcus lactis subsp. cremoris
            Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Lactococcus.
REFERENCE   1  (bases 1 to 1327)
  AUTHORS   Bolotin,A., Ehrlich,S.D. and Sorokin,A.
  TITLE     Studies of genomes of dairy bacteria Lactococcus lactis
  JOURNAL   Sci. Aliments. (2002) In press
```

COMMENT Contact: Sorokin A
  Genetique Microbienne
  INRA
  CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
  Tel: 33 1 34 65 25 16
  Fax: 33 1 34 65 25 21
  Email: sorokine@jouy.inra.fr
  best homologue in strain Il1403 is groEL (93%)
  Class: shotgun
  High quality sequence start: 30
  High quality sequence stop: 1299.

FEATURES Location/Qualifiers
 source 1..1327
  /organism="Lactococcus lactis subsp. cremoris"
  /strain="MG1363"
  /clone_lib="MG1363 Random Sequence Tag Library"
  /note="vector: pSGMU2; Site 1: SmaI; Strain Library of
  chromosomal fragments of L.lactis strain MG1363 was
  prepared by partial AluI digestion or by sonication."
  /db_xref="taxon:1359"

BASE COUNT 392 a 292 c 239 g 403 t 1 others
ORIGIN

Query Match 25.3%; Score 414.2; DB 17; Length 1327;
Best Local Similarity 72.0%; Pred. No. 1.7e-104;
Matches 553; Conservative 0; Mismatches 214; Indels 1; Gaps 1;

Qy 816 TGCTGTCAAAACGCCAGGAGATTTGGTGGTATGGTCTGTAAAGGTATGCTTGAGACAATTGCTAT 875
  |||||||| |||| | | |||| |||||| || ||| |||||||||||||| ||||| ||||||||
Db 1322 TGCTGTAAAAGCACCGGGATTTGGTG-CCCGCGTAAAGCTCAATTGGAAGATTTAGCCTAT 1264

Qy 876 CTTGACAGGTGGTACAGTGATTACACAGAGGATCTAGAGACTTGAATTAAAAGATGCTACAAT 935
  |||||| |||||||| ||| |||| || |||| | ||| | ||| |||| ||||| ||| ||||
Db 1263 CTTGACAGGTGGTACAGTAATACTGGAGGAACTTGGCTTTGATCTTAAAGATGCTACACT 1204

Qy 936 GACAGCCCTTGGACAGGCTGCTTAAGATTTACAGTTGATAAAGATAGCACACAGTAATTGTTGA 995
  ||| ||||||||||||||| ||||| |||| |||||| | |||| |||||||||||||||||||
Db 1203 AGAAGCTTTAGGACAAGCTGCTAAAGCATCTGTTGATAAAGAGCCATACAACAATTGTTGA 1144

Qy 996 AGGTTCAGGAGTTCAGAGAGTTCATTGCTAACCGTATTGCACTGATTAAATCGCAATTAGA 1055
  ||||||| ||||| | | |||| ||| |||| | | |||||||| ||| || ||||||| ||
Db 1143 AGGTGCTCGGTTCTGTGGTGCCATTTCTGATCGTGGCTATCATCAAAGCACAAAATCGA 1084

Qy 1056 AACAACAACTTCTGACTTTGACCGTGAAAAAACTACAAGAACGTTTGGCGAAATTAGCTGG 1115
  ||| |||| || | |||| | |||||||||| | ||| ||||||||||||||||||||||||
Db 1083 AAAAATAACTTCTGATTTTGCCCGTGAAAAATGNCACAGACGGCTTACGCGAAATTAGCTGG 1024

Qy 1116 TGGTGTGACCTGTTATCAAAGTAGGAGGCTCCACACAGACAGCTTTAAAAGAAATGAAACT 1175
  | ||| || |||| | || |||| | | | | |||||| || |||||| |||| |||||||
Db 1023 CGGAGTTCCGGTAGTTAAAGTAGGAGACCCGCACAGAACAGAACACTTAAAGCAATGAATCT 964

Qy 1176 TCGGATTGAGGATGCTCTAAAATGCTACACGTCGCAGCCGGTTGAAGAAGGTATCGTTGCTGG 1235
  |||| ||||| ||| ||||| | | | || || |||| || ||| |||||| |||||| |||
Db 963 TTTGATTGAAGACGCCACTAATGCAACACGTGCTTCCAGAGATGGCGGTATCGTTGCTGG 904

Qy 1236 TGGTGGAACAGCACTTATTACGGTTATTGAAAAATAGTAGCAGTCTTGAGCGTGAGGGCGA 1295
  ||||| ||||||| ||||||| | || || |||| | ||| || |||||| ||| ||| |
Db 903 TGGTGGTACAGCTCTTGTTAACGCCATCGCAGCTTTGGATAAACTTTCAGAAGAGGGCA 844

Qy 1296 TGATGCTACTGGACGTAACATTGTGCTTGTGTCGTCTTACGTCAAGTTGC 1355
  | | ||| |||| ||||| | |||||| |||||| |||||||||||||||
Db 843 CATTCAAACAGGAATTAACACGTCGTCGCGTGCCCTTGAAGACCCAGTTCGTCAAATCGC 784

Qy 1356 TTTAAATGCTGGGTACGGAAGGCTCCGTAGTTATTGACAAGTTGAAAAAACGCCCTGCAGG 1415
  | | || || |||| ||| ||||| || || |||||| || || ||||||||| ||||||
Db 783 GGGCCAATGCAGGATATGAAGGTTCAGTTNATCATTGGATTAAAACTTCGTTCAGAAAAGTTGG 724

Qy 1416 AACAGGATTTAATGCTGCCAACAGGTGAGTGGGGTTGATATGCAATTCATTGA 1475
  ||||| ||||||||| ||||| | |||| | |||| || || | ||| | |||
Db 723 TACAGGATTTAATGCGACGGACAGGTCAATGGGTAAATAATGATGAGAGGAATCGTTGA 664

Qy 1476 CCCTGTCAAAGTAACACGAATCAGCCGCTTCAAAAATGCAGCTTCTCGTAGCTAGTCTTATTTT 1535
  |||| |||| || ||| | | ||| | | ||||||| | | |||| | | || |||
Db 603 GACTACTCAACCAGTTGTCGGCTAATAAAACCAGAACCAGCTTGCTCCAGC 556

Db 663 CCCGGCAAAAGTGACTCGTTCAGCTTTGCAAAATGCAGAGCTTCTCGTGGCTGGATTAATCTT 604
  |||| ||||| | || || ||||||| | |||| |||| ||||||| | ||| || || ||
Qy 1536 GACAACAGAAGCAGTTGTTGCTAATAAAACCTGAACCAGCCTACGGCAGC 1583

Db 603 GACTACTCAACCAGTTGTCGGCTAATAAAACCAGAACCAGCTTGCTCCAGC 556

RESULT 5

LOCUS AY108560 2143 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PCO140434 mRNA sequence.
ACCESSION AY108560
VERSION AY108560.1 GI:21211654
KEYWORDS HTC.
SOURCE Zea mays.
 ORGANISM Zea mays.
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
  clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2143)
 AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
  Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
 TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
  Overgo Probes
 JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 2143)
 AUTHORS Coe,E.C.
 TITLE Direct Submission
 JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
  Missouri, Columbia, MO 65211, USA
FEATURES Location/Qualifiers
 source 1..2143
  /organism="Zea mays"
  /db_xref="MaizeDB:638266"
  /db_xref="taxon:4577"
  /clone="PCO140434"
  /clone_lib="Maize Mapping Project/DuPont Cornsensus
  Library"
  /note="this sequence is part of a project of EST
  assemblies resulting from the application of public
  contigs to seed DuPont contigs; this resource was
  assembled by DuPont as part of a collaboration for the
  overgo addressing of BACs in conjunction with the Maize
  Mapping Project"

BASE COUNT 546 a 494 c 598 g 505 t
ORIGIN

Query Match 23.4%; Score 383.6; DB 11; Length 2143;
Best Local Similarity 54.0%; Pred. No. 8.6e-96;
Matches 880; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

Qy 4 GCAAAAGAAATCAAAATTTTCAGCAGAATGCGCGTTGCTGCCCATGGTGCGGCGGAGTTGAFATG 63
  || ||||| || || |||| | | || ||| || |||| | ||| || || |||
Db 274 GCCAAGGAGAATCGCCTTTGACCAGGGCGTCCAGAGCGCGCCCTTCAGCGCCGGTCGAGAAG 333

Qy 64 TTAGCAGATACCGTCAAAGTAACCGCTTGGTCCTAAAGGCGCCAAATGTTGTTCTTGAAAAA 123
  | || ||| || | |||| ||| || | || | ||||| || | ||| | |||
Db 334 CTCGCCCCCCCGGGTCGGGCGGTCCACCCTCGGACCCAGGGAAGGAAATGTTGTCTTGGATGAG 393

Qy 124 GCTTTTGGTTCTCTCCCTTNATTACTNATTGACGGGTAACCATTGCTAAAGAGATCGAATTA 183
  ||||| | ||| | | | | | ||||| | ||| |||| | |||| | |||||
Db 394 ---TTTGGAACCCCAAAGTCGTCAAATGATGGAGTTACCATTGCCCGCGGCTATTGACGTT 450

Qy 184 GAAGACATTTTGAAAACATGGGAGCAAAAAATTGGGTGTCTGAAGTGGCTTCTAAAACCAAT 243
  | |||| || | |||| || | | ||| | |||| | ||| | | |||| |||||
Db 451 GCTGATCCCATGGAGAATGCTGGTGCTTCCTTGATTCGTGAAGTTGCTAGCAGACGGAT 510

Qy 244 GAPATTGCTGGTGATGGGACGAC-TACTGCACACAGTTTTGACACAGCCATTGTTCATGA 302
  || ||||||| | | | |||| | | | ||| || | |||| | | | ||| |
Db 511 GACTCGGCTCGGTGATGGAACCACAAAACTGCCCTCTCGTTCTCCGGCTAGGGAGAGTCAATCAAAATT 570

Qy 303 AGGACTAAAAAAATGTGACACGCAGGTGCTAATCCAATTGGTATCGTCTGCGAGGCATTGAAAC 362
  || | | | | | | ||||| || | ||||| | ||||||| |
Db 362

```
Db   571  GGGCATGGTTGAGCATTACTTCAGGGGCAAATCCAGTGGTCCGTTAAGAAGGGCATTGATA  630
Qy   363  AGCAACGCAACAGCTGTTGAAGCCTTGAAAGCCATTGGCTCAACCTGTATCTGGCAAGGA   422
Db   631  GACGTTCAGAAATTGGTGGAGGAACTTGAGAGAAATCAAGGCCGGTTCAAGGGTAGTGG    690
Qy   423  AGCTATTGCTCAGGTCGTCGCTGACTAATCATCACGCTCTGAAAAAGTTGGAGAGTAT---AT  479
Db   691  GGATATTAAAGCTGTTGCTGCTGCCATATCAGCTGGGAAATGAATTGTTGTGGGACTATGAT  750
Qy   480  CTCAGAAGCTATGGAGACGTGTGGGCAACGATGGTGTGATTACCATCGAAGAAATCTCGAGG  539
Db   751  CGCCGAAGCTATTGACAAGGTTGGGCCCTGATGGTGTCCTCTCAATTGAGTCGTCATCGTC  810
Qy   540  TATGGAAACAAGAACTTGAAGTGGTTGAAGGCATGCAATTGACCGTGTGGTTACCTGTCTCA  599
Db   811  GTTTGAGCACACAGTTTGAAGTTGAAGAAGGGATGGAGCTTGACGAGGATAATACTCCCC   870
Qy   600  ATACATGGTCACACAGACAATGAAAAATGGTTGCCAGACCTTGAAAACCCCATTTATCTTAAT  659
Db   871  TCAGTTTGTCACTAACCTTGAAAAATCTATTGTTGACTTGGAAATGCTCGGATTCTTGT   930
Qy   660  CACGGATAAAAAAGTTGTCAAACATCCAAGACANTTTGCCACTACTTCAGGAAGTTCTTAA  719
Db   931  CACTGATCAGAAGATATCATCGGATAAAAGAAAATTATTCCTCGTTGGAGCAGAACAAACA  990
Qy   720  AACCAACCGTCCATTACTTCATTATTGCAGATGATGTGGGAGAGGGTGAAACCACTTCCAAACCCT  779
Db   991  GTTAGAGAGCACACTTCTTTATAATTGCGAGGATGTAAGTGCCGGAGGGCGCTGGCCAACATT  1050
Qy   780  TGTCTTGAACAAGAATTCGTGGTGGTACTTTCAATGTGGTTGCTGTCTCAAACGGCCAGGATTTGG  839
Db  1051  AGTCATAAACAAGCTTAGAGGAATTTTAAATGTGCCGTCCAATCAAAGCTCCTGGTTTTGG  1110
Qy   840  TGATCGTCGGTAAGCTAAGCTATGCTTGAAGACATTGCTATCTTGACAGTGGTACAGTGATTAC  899
Db  1111  TGAGAGCGTAAAGCTCTTCTTCAGGACAATTGCCATTGTTACAGGTGCTGAAATATCAATC  1170
Qy   900  AGAGGATCTAGGACTTGAAATAAAAGATGCTACAATGACAGCCCTTGGGACAGGCTGCTAA  959
Db  1171  CAAAGATCTTGGTGCTTACTAGTTGAGGATACACACGTGGCAGCTTGGCATAGCTCGGAA   1230
Qy   960  GATTACAGTTGATAAAGATACACAGTAATTGTTGTTGAAGGTTCAGGAAGTTCAGAAGCTAT  1019
Db  1231  AGTTACAATCTCAAGTTCCTCGACGACCATTATTAGGCAGTGCTGCTAGCAAAGACGATAT  1290
Qy  1020  TGCTAACCGTATTCGCACTGATTGATTAAATCGCAATTAGAAAACAAACTTCTGACGTTGACCG  1079
Db  1291  CCAGGCCAGAATCGGCCGAGCTGAAGAGAGCTTTCTCAGACGGACTCGACGTATGACTC   1350
Qy  1080  TGAAAAACTACACAGAAACCTTTGGCGAAATTAGCTGGTGCTGTTATCAAAGTAGG     1139
Db  1351  TGAGAAGTTGGCAGAGAATTGCAAAGCTTTCTGGTGAGTTGCTCGTTGGTTAAGGTTGG   1410
Qy  1140  AGGTCCAACGAGAGACAGCTTTAAAAGAAAATGAAACTTCCCATTGAGGATGCCTCTAAATGC  1199
Db  1411  AGCATGACCGAGGCACAGAGCTAGGAGGACCGAAGCTCCCCATAGAGGACCGGAAGAATGC  1470
Qy  1200  TACACGGTGCAGCCGTTGAAGAAAGGTATCGTTGCTGGTGGTGGAACAGCACTATTACGGT  1259
Db  1471  GACTTTTGCAGCAATAGGAGGAAGGTATTGTTCAGAGGCTGCACAGATTGTTCATCT     1530
Qy  1260  TATTGAAAAGTAGCACCTCTTGACTTGAGCGGCGATGATGCTACTGGACG-------      1310
Db  1531  ATCGACACATTCGTACCGGCTATCAAGGACGGTTGGATGATCCTGAGGGAGGCCCTTAGGTGC  1590
Qy  1311  TAACATTGTGCTTCGTCGTCTCTGAGAGACGCGTAGGTATGGAAACAGAACAAGCCATTGGGTA  1370
Db  1591  TGATATCATTCAGAAGCCTTAGGTGGGCACCTGCAGCCTGATGAGCCCATAATGCTGGGCT   1650
Qy  1371  CGAAGGCTCCGTAGTTATTGACAAGTTGAAAAAACACCCTGCAGGAACAGAGATTTAATGC  1430
Db  1651  GGAAGGTGAGGTGATTGTTGGATAAAAATCCGGGAAAAGCGAGTGGGAGTTCGGTTACAAGCG  1710
```

RESULT 6
AY106465

```
LOCUS       AY106465               1376 bp    mRNA       linear   HTC 25-MAY-2002
DEFINITION  Zea mays PCO128697 mRNA sequence.
ACCESSION   AY106465
VERSION     AY106465.1  GI:21209543
KEYWORDS    HTC.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 1376)
  AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE     Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
  JOURNAL   Unpublished (2002)
REFERENCE   2 (bases 1 to 1376)
  AUTHORS   Coe,E.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1..1376
                     /organism="Zea mays"
                     /db_xref="MaizeDB:637486"
                     /clone="PCO128697"
                     /clone_lib="Maize Mapping Project/DuPont Consensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      397 a    277 c    359 g    343 t
ORIGIN

Query Match         18.5%;  Score 302.8; DB 11; Length 1376;
Best Local Similarity   55.7%;  Pred. No. 2.8e-73;
Matches   621;  Conservative    0;  Mismatches 487; Indels  6;  Gaps  2;

Qy  1431  TGCAACAGGTGAGTGGGTTGGATATGATTAAAACAGGAATCATTGACCCTGTCAAAGTAAC  1490
Db  1711  CATGCGCGACAAGCACGAGAACCTGTGTGAGGCGTGGTTGGATGCGACCCTGCCAAAGTGAC  1770
Qy  1491  ACGATACGCGCTTCAAAATGCAGCTTCTGTTAGCTAGTTTATTTGACAACGAAGCAGT   1550
Db  1771  TAGGTGCGCCCCTGCAGAACGTGCCTCCCTGGCCTGGAATGGTGCTGACTACCAAGCAAT  1830
Qy  1551  TGTTGCTTAATAAACCTGAACCAGCTACGCCCACCGCCAGCAATGCCAGCCAGGTATGGATCC  1610
Db  1831  CGTTCTGGAGAAGGCCCACAGAAGGCTCCAGCAGCGAGCGTCGCCACGGCACCATAGGGTTC  1890
Qy  1611  AGGAATGATG  1620
Db  1891  TTTCACCATG  1900
```

Qy 464 AAGTTGGAGAGTATATCTCAGAAGCTATGGAGCGGTGTGGGCAACGATGGTGTGATTACCA 523
Db   23 AAATTGGAGAGATCTAATATCAAAAAGCCATTGGAAAAAGTTGGAAAAGATGGAGTCATTACTA 82
Qy 524 TCGAAGAANTCTCGAGGTATGGAAACAGAACTTGAAGTGGTTGGTCAAGGCATGCAATTTGACC 583
Db   83 TTGTTGATTGGCAAAAACANTTGGACANTAGCTTGAAGCACTACAGGGAATGAAGAGCTGTCCA 142
Qy 584 GTGGTTACCTGTCTCANTAACATGGTCACAGACAATGAAAAAATGGTTGCCAGACCTTGAAA 643
Db 143 GAGGATACATATCTCCTTACTTGTGACTGATCAAAAAGACTTCAGAAAATGTGAGATGGAGA 202

```
SOURCE       Zea mays.
  ORGANISM   Zea mays
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
             clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1  (bases 1 to 1387)
  AUTHORS    Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
             Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
  TITLE      Maize Mapping Project/DuPont Consensus Sequences for Design of
             Overgo Probes
  JOURNAL    Unpublished (2002)
REFERENCE    2  (bases 1 to 1387)
  AUTHORS    Coe,E.C.
  TITLE      Direct Submission
  JOURNAL    Submitted (25-APR-2002) Maize Mapping Project, University of
             Missouri, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1..1387
                     /organism="zea mays"
                     /db_xref="MaizeDB:637188"
                     /db_xref="taxon:4577"
                     /clone="PCO123570"
                     /clone_lib="Maize Mapping Project/DuPont Consensus
                     Library"
                     /note="this sequence is part of a project of EST
                     assemblies resulting from the application of public
                     contigs to seed DuPont contigs; this resource was
                     assembled by DuPont as part of a collaboration for the
                     overgo addressing of BACs in conjunction with the Maize
                     Mapping Project"
BASE COUNT      426 a    257 c    372 g    332 t
ORIGIN
```

```
Query Match        17.7%;    Score 290;  DB 11;    Length 1387;
Best Local Similarity  56.0%;  Pred. No. 1.1e-69;
Matches  593; Conservative    0; Mismatches  455;    Indels  10;   Gaps   2;

Qy    554 TTGAAGTGGTTGAAGGCATGCAAATTTGACCGTGGTTACCTGTCTCAATACATGGTCACAG 613
          ||||||| |||| ||| ||   ||||||| ||| |||  |  | ||| ||||   ||||||
Db     14 TTGAAGTTGAAGAAGATGGAATTGACGCGCGGTTATATTTCCCACAATTTGTGACAA  73

Qy    614 ACAATGAAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATCACGGATAAAAAAG 673
          |||| |   |  | || || || |||  || || || || | | || | ||||  ||||||
Db     74 ACCTCGAGAAATCAATTGTGGGAGTTTGAGAATTGTAAGGTTCTTATTACTGATCAGAAGA 133

Qy    674 TGTCAAACATCCAAGACATTTTGCCACTACTTCAGGAAGTTCTTAAAACCAACCGTCCAT 733
          |||  | ||  | || ||| |  | || ||||||||| |||| ||||| ||||| ||||||
Db    134 TCACAAGGCATAAAAGGAAAATCCTTCCGATTTGGAGAAGACCACACAGCTGAGCGCCCTC 193

Qy    734 TACTCATTATTCGCAGATGAGTGTGGATGGTCAAGCACTTCCAAACCCTTGTCTTGAACAAGA 793
          || |   | | |||| |||  || || || ||| ||| ||| ||| ||| || |||| ||||
Db    194 TGTTCATTATTGCTCGAGGACATTACTGGTCGAAGCTTTGGGCAAACTCTTGTTAACAAGC 253

Qy    794 TTCCTGTGGTACTTTCAAATGTGGTTGTCTGTCAAAACGGCCAGGATTTGGGTGATCGTCGTAAAG 853
          |||  | |||  | || || ||| || |  || ||||| || ||||  | || |||||  |||||
Db    254 TTCCAGGAAATTCTCAAATGTTGCAGGGATTAAGGCTCCAACGTTTGGGGTGAGCGGGGAAGG 313

Qy    854 CTATGCCTTGAAGCACATTGCTATCTTGACAGGTGGTACAGTGATTACAGAGGATCTAGGAC 913
          |||| || || |  ||| || | | ||| ||| || ||||||||  ||||||| ||| ||||
Db    314 CTGTACTTCAGGACATTGCCATCGTCACAGGTCCTGAGTTCCTAGCAAAAGATCTTGGTT 373

Qy    914 TTGAATTAAAAGAGTGCTACAATGACAGCCCCTTGGACAGGTGTCTAAGATTAGAGTTGATA 973
          || |  |||  |||||||| | | ||| || || ||  || || || | | | |||| |||
Db    374 TGTTGGTTGAAAATGCTACAGAGGAGCAACAACTGGGCACAGCAAGAAAGTCACGATACATC 433

Qy    974 AAGATAGCACACTAATTCTTGAAGGTTCAGGAAGTTCAGAAAGCTATTGCTAACCGTATTG 1033
          ||||| ||| || | || || || || || | ||  || || |||| | | ||| |||
Db    434 AGACTACAACAACCCTCATAGGAGATGCAGCCAGTAAAGATGAGATCCAGGCGAGGATTG 493

Qy   1034 CACTGATTAAAATCGCAATTAGAAACAACAACTTCTGACTTTGACCGTGAAAAAACTACAAG 1093
          ||| ||||| || | || || || | |||  ||||||  | || | || |||| | ||||
Db    494 CACAACTGAAGAAGGAGCGTTGCTGAGACTGAATCAGTTTATGATACTGAGAAAATTGGGCCG 553
```

```
Qy    644 ACCCATTTATCTTAATCACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTAC 703
          |||| | |||||||| ||| || || ||||| ||  || | || | |||| | |  |  |
Db    203 ACCTCGTATTACTTATCCATGACAAGAAAATCTCAAACATGGATTCTCCTTCCAGCAT 262

Qy    704 TTGAGGAAGTTCTTAAAACCAACCGTCCATTACTTCATTATTGCAGATGATGGTGGATGGTG 763
          ||| || |||  || |||| |||| ||||| |||| | | ||||| |||| |||| |||||
Db    263 TAGAAATTTCTATCAAGAATCGGAACCGTCCTTCTTCTCATTGTTGCTGAGGATGTTGAAGGAG 322

Qy    764 AAGCACTTCCACACCTTGTCTTGAACAAGATTCGTGGTGTACTTTCAATGTGGTTGCTGTCA 823
          || ||| || | ||  ||| ||| ||||||  | || || | || | | || || || ||
Db    323 ATGCTGTCAATGCTTGTACTGAACAAGACATCTGTGGACTCAAGGTATGTGCTGTCA 382

Qy    824 AAGCGCCAGGATTTGGTGATCGTCGTAAAGCTATGCCTTGAAGACATTGCTATCTTGACAG 883
          ||| |||||| |||| ||||| | ||||||| || || || ||| || | || | ||||
Db    383 AAGCTCCTGGATTTGGTGAAAATAGAAGGCACAATCTAGATGACACATGGCTTGTGACTG 442

Qy    884 GTGGTACAGTGATTACAGAGGATGGTCTAGGACTAGGACTTGAAATTAAAGATGCTACAAGACAGCCC 943
          |||| || ||||| | |||  | ||   ||| | |||| | |  || | || || | ||||||||||
Db    443 GAGGAGAGGTTATTAGGGAGGAACGGGGTCTTGATCTTGATCTTGAAATTACAAATGC 502

Qy    944 TTGGACAGGCTGCTAAGATTACAGTTGATAAAGATAGCACAGTAATTGTTGAAGGTTCAG 1003
          || | |||||| |||| | | | | ||| ||||  | || ||||| |||| |||| ||||
Db    503 TTGGCACTGCTAAAAAGGTAACTGTATCCCTTGATGATACTATCATCCTAGATGATGGGAG 562

Qy   1004 GAAGTTCAGAAGCTATTGCTAACCGTATTGCCACTGATTAAAATCGCAATTAGAAACAACA 1063
          ||||||| || | |||| | |   ||| | || ||| ||||||| | ||||  ||| || |
Db    563 GTGACAAACAGCAGATAGATGAGATGTGCCAACAGCTTAGAATCAATTGACACAAGCA 622

Qy   1064 CTTCTGACTTTGACCGTGAAAAAACTACAAGAACGTTTGGCGAAAATTAGCTGGTGGTGTAG 1123
          |||| ||||  || | || |||| |  | || |||  || | ||||| | || || ||  |
Db    623 CTGCTGTATTCGACAAAGAAAAGGCCCAAGAGGCTTTTATCTAAGTTATCTGAAGGTGTTG 682

Qy   1124 CTGTTATCAAAGTAGGAGCTCCAACAGCAGACAGCTCTTTAAAAGAAATGAAAACTCCCAATG 1183
          |||  || |   || || || ||| || || | |||  | ||| |  | ||| || ||||||
Db    683 CTGTACTCAAGATTGGTGGAGCTAGTGAAGCTGAAGTTGGTGGAGAAGAAAGATAGGGTGA 742

Qy   1184 AGGATGCCTCTAAATGCTCACACTGCAGCCGTTCAGCACCGTATCGTTGCTCGGTGGTGGAA 1243
          ||||| ||||  |||| || || || ||| |||| || | | ||   | |||||  || |||
Db    743 CAGATGCTCTAAATGCTGCAAGGCTTCTCGTGGAGGAGGCATTGGTGCCAGTGGTGGGAG 802

Qy   1244 CAGCACTTATTACGGTTATTGAAAAAGTAGCAGCTCTTGAACCTCTTGGAGGCGAGTGGTATCGTTGCTTGACCTTG 1300
          ||||  || || | |||||| | |  | || || ||  || ||  || | |  | || || | || || |||
Db    803 TTGCCCTTCTTTATGCCACCAAGGAGCTTGACAAGGATTAGGCACAGCAAAACGAGGACGAAA 862

Qy   1301 CTACTGGACGTAACATTGTGCTTCGCGTCGTGGTCTAGGACTACGTCAAATTGCTTTTAA 1360
          ||  | ||| ||  || | | || | | | ||| || ||| | ||   ||||| ||  ||
Db    863 AGAATTGGAGTTCAAAATTATCAAGAATTCTTGAAGGCTCCCTTGATGACCATAGCTGCAA 922

Qy   1361 ATGCTGGGTACGAAGGCTCCGTAGTTATTGACAAGTTGA---AAAACAGCCCTGCAGGAA 1417
          ||| | |  || |  || | || | || ||||||  |||   |||| | | |||||||
Db    923 ATGCTGGATCATTGGAGTGCTATAGTCATAGGGAAGTTGGCATAGGAAGCGGACCTCAGTT 982

Qy   1418 CAGGATTGAAATGCTCGCACAGGTGAGTGGGTTGCATAGTGAATGATTAAAACAGGAATCATTGACC 1477
          ||||| || |  | | | || || || || |||  || || ||| |||| ||||   | |||||
Db    983 TGGGCTATGATGCAGCAAAAAGGGGAGTAGGTAGCAACATGATCAAGAGGCATCATCGATC 1042

Qy   1478 CTGTCAAAGTAACACGATCAGCGCCTTCAAAAATGCACAGCTTCTGTAGCGTAGTCGTTATTTGA 1537
          ||| |  ||| | ||| | ||  | || |||  | |||| ||  || | || | | | || ||
Db   1043 CCGTGAAGGTGAATCCACCGCACTTCAACAGCGCTGCAAGTGTTCCCTGCTGATGGGCAA 1102

Qy   1538 CAACAGAAGCACGTTGGTTGCTAATAAACCTGACC 1571
          ||| || | ||| || || | |||| |||| ||
Db   1103 CCACGGAGGAGGCCGGCTGTCTCCGAGCTCCCGGCCAAC 1136
```

```
RESULT 7
LOCUS       AY108518            1387 bp    mRNA    linear   HTC 25-MAY-2002
DEFINITION  Zea mays PCO123570 mRNA sequence.
ACCESSION   AY108518
VERSION     AY108518.1  GI:21211596
KEYWORDS    HTC.
```

```
Qy  1094 AACGTTTGCGAAATTAGCTGGTGGTGGTGTAGCTGTTGTTATCAAAGTAGGAGGCTCCAACAGAGA 1153
         |  |  ||  |  ||| | |||||||||| || ||||| ||||| || ||| |||||   || ||
Db   554 AGAGAATTGCTAAGCTTGCCGGTGGTGGTTGTTGCCGTGGTTATCAAGGTGGAACTGCAACTGAGA 613

Qy  1154 CAGCTTTAAAAGAAATGAAAACTTCGCATTGAGGATTGCTCTAAAATGCTTACACGTGCAGCCG 1213
         | || ||| ||| || |   ||| |  |||||| |||||| |||||||| ||| |||||||| |
Db   614 CTGAGCCTTGAGGACCGGCAGCTAAGAAATTGAGGAATGCCAAAAAATGCCACTTTTGCAGCTA 673

Qy  1214 TTGAAGAAGGTATCGTTGCTGGTGGTTGGTGGTGGAACAGCA----------CTTATTACGGTTATTG 1264
         | |||| | | | |||||| ||||||| |   | ||| |||          ||||| ||||| ||
Db   674 TAGAGAGGGCATTGTTCCAGGAGGTGGTCCGGCATATGTCCACCTATCTACGGTTTGTCC 733

Qy  1265 AAAAAGTAGCACCTCTTGAGCTTTGAGGCCGATCATGATGCTACTGGGACGTRACACATTGTGCTTC 1324
         ||   || || |||| |||| || | |  ||| ||||||||||| |||  || ||||  |  | ||
Db   734 CTAAGATCAAAGAGACCAATTGAGGAGACCCTTGGGCGGCCTTGATTAATTATTCAAA 793

Qy  1325 GTGCTCTAGAAGAGAGCCCTGTACGTCAAAATTGCTTTAAAATGCTTGGGTACGAAGGCTCCGTAG 1384
         ||||| || || || ||||| || || ||||||||| |||||||||||| ||| |||||||| ||
Db   794 AGGCAATTGGTAGCACCTGCTTCATTGATTCCACACAATTGCTGGACTAGAGGGTGAAGTGG 853

Qy  1385 TTATTGACAAGTTGAA-AAACAGCCCTGCAGGAACAGCAGATTTAATGCTGCAACAGTGAG 1443
         |||| ||||||||||  |||||| ||||| |||||| ||||| |||||| ||||| ||||
Db   854 TTGTCGAGAAGAATCAAAGGCAGGAGAGCCGATTGGGGAGGTGGCTACAATGCAAATGACTGACAAG 913

Qy  1444 TGGGTTGATATGCATTAAAAACAGGAATCATTGACCCTGTCAAAGTAAACACGAACAGCGCTT 1503
         ||| ||||| |  || ||||||| |||| ||||| |||| |||||| || ||||||||| ||
Db   914 TACCAGAAACCTGATGGAGGACCCTGGTGGTTGATCGAGCCGCCAAGGTGACCAGGTGTGCGCTC 973

Qy  1504 CAAAATCGCAGCTTCTGTGTAGCTAGTCTTATTTGAACACAGAAGCAGTGTTGTTCGTCAATAAA 1563
         ||||| ||| ||| ||| ||  ||| | | |||| |||| | ||||| ||||||| | || |||
Db   974 CAGAATGCGGCGGCCTCAGTAACCTGGAATGGTCCTAACCACGGCAGGCCATTGTTGTTGAGAAG 1033

Qy  1564 CCTGAACCAGGTACGCCCACCGCCACCAAATGCCAGCAGG 1601
         || || |||| | ||||||||||||||||||| ||||||
Db  1034 CCAAAGCCGAAGCCTCAGAGTGGCACGAGAGCCACCGGAGG 1071
```

RESULT 8

```
LOCUS       CNS074DM                 887 bp    DNA     linear   GSS 07-JUL-2001
DEFINITION  Kluyveromyces lactis, genomic survey sequence.
ACCESSION   AL428672
VERSION     AL428672.1  GI:12211866
KEYWORDS    GSS.
SOURCE      Kluyveromyces lactis.
  ORGANISM  Kluyveromyces lactis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
REFERENCE   1  (bases 1 to 887)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 887)
  AUTHORS   Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F.,
            Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R.,
            Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 11.
            Kluyveromyces lactis
  JOURNAL   FEBS Lett. 487 (1), 66-70 (2000)
  MEDLINE   20584721
   PUBMED   11152886
REFERENCE   3  (bases 1 to 887)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1..887
                     /organism="Kluyveromyces lactis"
                     /strain="CLIB 210"
                     /variety="lactis"
                     /db_xref="taxon:28985"
                     /clone="BA0AB027B07"
                     /clone_lib="BA0AB"
     misc_feature    <4..>887
                     /note="similar to Saccharomyces cerevisiae ORF YLR259c [
                     HSP60 ; heat shock protein ' chaperone, mitochondrial ]"
                     /evidence=not_experimental
BASE COUNT      256 a    163 c    216 g    252 t
ORIGIN
```

```
Query Match    16.1%;  Score 263.2;  DB 17;  Length 887;
Best Local Similarity 58.7%;  Pred. No. 2.7e-62;
Matches  493;  Conservative   0;  Mismatches 343;  Indels   4;  Gaps   2;

Qy     7 AAAGAAATCAAATTTTCAGCACGATACGCCGTTGGTCCTAAAGGCGCAATGTTGTTCTTGAAAAGCT 126
         |||||||| ||| ||||  || || | |||| ||| | | |||  | | || ||| | | ||||| | |
Db    25 AAAGAAATTAAAGTTTGGTGTTGAAGGTAGAGAGCTGCCTTGTTTAAAGGGTGTTGAAAACTCTA 84

Qy    67 GCAGATACCGTCAAAGTAACGCTTGGTCCTAAAGGCGCAATGTTGTTCTTGAAAAGCT 126
         || || ||| ||  |  ||||  || || | || |  || ||||| |||| |||| |
Db    85 GCGGAGCTGTTTCTGCCACTTTGGGTCCCAAAAGGTAGAAAACGTTTTGAATTGAGCAACACCA 144

Qy   127 TTTGGTTCTCCCTTAATTACTAATCACGGGTAACCATTGCCTAAAGAGATCGAATTAGAA 186
         ||||  | ||||  ||| || |  |||| || | | ||||| |||||   | ||| | ||
Db   145 TTTGCGCCTCCAAAGATCACCCAAAAGGATGGTTACTGTTGCTAGAGCCATTACTTTGGAA 204

Qy   187 GATCATTTTGAAAAACATGGGAGCACAAAATTGGGTGTCTGAAGTAGCCTCTAAAACCAATGAT 246
         |||  |||||||||||||||| || | ||  |||   | |||| ||||| | || ||   |||
Db   205 GATAAGTTCGAAAAACATGGGTGCCTAAGCCTTTTACAAGAAGTTGCTTCCAAGACAAAACGAA 264

Qy   247 ATTGCTGTGGTGATGGGACCGACTACTGCAACAGTTTTGACACACAGCCCATTGTTCAGAAGGA 306
         | |||  | || | ||| || ||||| | || ||| | |  | | |||| |||| |  ||| |
Db   265 GCTGCCGGTGACCGGTACCACTTCGGCCACTGTGTTCTTAGGTAAGGCCATTTCACTGAAATCC 324

Qy   307 CTAAAAAATGTGACAGCAGGGTGCTAATCCAATTGGTGTTATCCGTCGAGCGCAATTGAAACAGCA 366
         || ||||||||| || | || ||| | | ||| ||||| ||||   |||| ||||||| |||| |
Db   325 GTTAAGAATGTTGCAGCTGTGGTTGTTGTTAACCCAATGGATTTGAGAAGAGGTACTCAAGCTGCT 384

Qy   367 ACAGCACAACAGCTGTTGAAGCCTTGAAAAGCCATTGCTCAACCTGTATCTGGCAAGGAAGCT 426
         | || |||  |  ||||||||   | |||||  ||| ||| |   | || ||||| | || |
Db   385 GTCGAAAAAGTAAATTCAATACTTGTCAGGAGACAGAAAAGAAAATCACCACTTCTGCTACA 444

Qy   427 ATTGCTCAGGTCGGTCAGTATCATC---ACGCTCTGAAAAAGTTGGAGGTATATCTCA 483
         || ||||| || ||| || || |||   || || | |||||||||  |||| |  | |
Db   445 ATCGCTCAAGTGGCAACTATTTTCTGCCAAACGGGGATGCCTCAGCGTTGGTAAGGTTGTTGGCT 504

Qy   484 GAAGCTATGGAGGCGGCGGTTGGGTCGGGGCAAACGAGATGGTGTTGTGATTGACTACCATGG 543
         ||||| |||||| || || || || |  | ||||  || |||| | | || | ||| ||  |
Db   505 TCAGCCCATTGGAGAAAAAGTTGGGTAAGGAAGGTGTCATCACTATCACGATACAGAGACCTTG 564

Qy   544 GAAACAGAACTTGAAGTGGTTGAAGGCATGCAAATTTGACGGTGGTTACCTTGTCTCAATAC 603
         ||| | ||| |||||||  | |||||  ||| || || | |||||| | |||||| |||||
Db   565 GAGAGTGAATTGGAAGTCACTGAAGGTATGAGAATTCGACCGTGGTTTCATTTCTCCATAC 624

Qy   604 ATGGTCACAGACAATGAAAAAAATGGTTGCAGACCTTGAAAAACCCATTTAACTTAATCACG 663
```

```
Qy    625 TTTATTACCGACGCTAAATCTGGTAAGGTGGAATTTGAAAAACCTTGATCTTATTGAGT 684
              | ||    | ||  || || || |||||| |  ||||| |  ||||| | |  |
Db    664 GATAAAAAGTGTCAAACATCCCAAGACATTTTGCCACTACTTGAGAAGTTCTTAAACC 723
              |||| |   | ||||
Db    685 GAAAAGAAGATCTCCTCAAATCCAGATAATCTGCCTTCTTTGAAATTGTCAAAACAAACC 744
Qy    724 AACCGTCCATTACTCCATTATTGCCAGATGATGATGGCACTCCCAACCCTTGTC 783
              |||||| |||||||| ||||||| | || |||| || |||| |  ||||||
Db    745 AGAAGACCATTGTTGATCATCCCCGAAGATATTGATGGGG-ACTTTTGCTTGTATT 803
              |||||||||||| ||| ||| | | ||| ||||||| ||| |||||| ||||||
Qy    784 TTGAACAAGATTCGTTGGTACTTTCAATGTGGTTGCTCAAACGCGCCCAGGATTTGGTGAT 843
              |||||| ||| | ||| || |||||||||||| ||  |||||  ||||  | |||| ||
Db    804 TTGAACAAGTTGAGAGGTCAAGTTCAAGTTTGTGCGGTTAAAACTCCAGGTTTCGGTGAT 863
```

RESULT 9

```
LOCUS           BM799922
DEFINITION      BM799922            1090 bp    mRNA    linear    EST 05-MAR-2002
                AGENCOURT_6417164 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5492101
                5', mRNA sequence.
ACCESSION       BM799922
VERSION         BM799922.1  GI:19116745
KEYWORDS        EST.
SOURCE          human.
  ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1  (bases 1 to 1090)
  AUTHORS       NIH-MGC http://mgc.nci.nih.gov/.
  TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL       Unpublished (1999)
COMMENT         Contact: Robert Strausberg, Ph.D.
                Email: cgaps-r@mail.nih.gov
                Tissue Procurement: ATCC
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLAM12113  row: k  column: 14
                High quality sequence stop: 743.
FEATURES             Location/Qualifiers
     source          1..1090
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5492101"
                     /clone_lib="NIH_MGC_67"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.75 kb. Library constructed by Life
                     Technologies."
BASE COUNT      332 a    209 c    274 g    275 t
ORIGIN
```

```
Query Match     16.0%;  Score 262.8;  DB 14;  Length 1090;
Best Local Similarity  56.2%;  Pred. No. 4e-62;
Matches  514;  Conservative   0;  Mismatches 397;  Indels   3;  Gaps   1;
```

```
Qy      4 GCAAAAGAAATCAAATTTTCAGCAGATCGGCGTCGCCCATCGGTGGCGCGGAGTTGATATG 63
            || ||||| ||||||| |||||||  ||||||||||||  || |||||  || ||||
Db    122 GCCAAAGAGTGTAAAATTGGTGCAGATGCCCGAGCCTTAATCGCTTCAAGGTGTAGACCTT 181
Qy     64 TTAGCAGATACCGTCAAAGTAACCCCTTGGTCCTAAAGGCGCAAGTGTTGTTCTTGAAAA 123
            |||||||  |||||||| || || |||| ||| |||| |||||||||| | |||||
Db    182 TTAGCCGATGCTGTGGGCCGTTACAAATGGGGCCAAAGGAAGAAACAGTGATTATTGAGCAG 241
Qy    124 GCTTTTGGTTCTCTCCCTTAATTACTTAATGACGACGGGGTAACCATTGCTAAAGAGATCGAATTA 183
            | ||| ||| | |  |  |||| |||| ||| | |||||||  || || ||||| |||| ||||
Db    242 AGTTGGGGAAGTCCCAAAGTAACAAAAGTGTGTGCAAAGTCAATTGACTTA 301
```

RESULT 10

```
LOCUS           BF275584
DEFINITION      BF275584            878 bp    mRNA    linear    EST 07-MAR-2001
                GA__Eb0024G23f Gossypium arboreum 7-10 dpa fiber library Gossypium
                arboreum cDNA clone GA__Eb0024G23f, mRNA sequence.
ACCESSION       BF275584
VERSION         BF275584.1  GI:11206654
KEYWORDS        EST.
SOURCE          Gossypium arboreum.
  ORGANISM      Gossypium arboreum
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE       1  (bases 1 to 878)
  AUTHORS       Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
                ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE         An integrated analysis of the genetics, development, and evolution
                of the cotton fiber
  JOURNAL       Unpublished (2000)
```

```
Qy    184 GAAGACCATTTTGAAAACATGGGAGGCAAAATTGGGTGTCGAAGTGGCTTCTAAAACCAAT 243
            || |||  | || ||||| ||| || |||  ||||| | ||| ||| || | || |||||
Db    302 AAAGATAAAAATACAAGAGACATTTTGGAGCTAAACTTGTTCAAGATGTTCCCAATAACAAAT 361
Qy    244 GATATGTGCTGGTGATGGGACGGACTACGCCCAACACAGTTTTGACACAAGCCATTGTTCATGAA 303
            ||| |  ||   | ||| || |  || || |  |||  ||||| ||  |||| | | |||||
Db    362 GAAGAAGCTGGGAGATGGCGACTACCACTGCTACTGTGTACTGGCACGTCTATAGCCACAAGGAA 421
Qy    304 GGACTAAAAATGTGACAGCAGGGTGCTAATCCAATTGGGTATCCGTCGAGGCCATTGAAACA 363
            ||  ||||||  |||| |||||  || ||| ||||||| |  || || |||| ||| |||||
Db    422 GGCTTCGAGAAGATTAGGCAAAAGGTGCCTAATCCAGTGGAAAATCAGGAGAGAGGTTGATGTTA 481
Qy    364 GCAACAGCAACACACTGTTGAAGGCCTTGAAAAGCCATTGCTCAACCCTTATCTGGCAAGGAA 423
            || |  |  |  | |||| ||||  | || |||| |||| |||||  |  ||| || |||||
Db    422 GGCTTCGAGAAGATTAGGCAAAAGGTGCCTAATCCAGTGGAAAATCAGGAGAGAGGTTGATGTTA 481
Qy    424 GCTATTCCTCAGGTCCGTCGCAGTATCATCACCCGCTCTGAAAAAG---TTGAGACTATATC 480
            || |  ||||  |  | || || ||||| || || ||| || ||||     || ||| ||||
Db    542 GAAATTGCACAGGTTGCTACGGTTGCTAGGAAGGTCAACGAGGCTATATTGGCAATATCATC 601
Qy    481 TCAGAAGCTATGGAGGCCGTGTCGGGACCGATGGTGTGATTTACCATCGAAAGAAATTCGGAGGT 540
            |||| | ||||| || |||| ||||| || ||  ||| | || |||  | | |  ||| |||||
Db    542 GAAATTGCACAGGTTGCTACGGTTGCTAGGAAGGTCAACGAGGCTATATTGGCAATATCATC 601
Qy    541 ATGGAAACAGAACTGAGAGTGGTTGAAGGCCATGCAAATTGACCGTGGTGGTTACCTGTCTGCACCAA 600
            |||| |  | || || || ||| || || |||   ||||| || | || || |  || | ||
Db    662 CTGAATCGATCGAAATTGACAAAATTATTGTAGCGCATGAAGTTTGAGTGAGGCTATGGTTTCCCA 721
Qy    601 TACATGGTCACACGACAATAGAAAAAATGGTTGCAGATGGTTGATCAGGACCTTGAAAACCCATTATCTTAATC 660
            || || ||  |  | |||| ||||| |||| || |  || | || |||| || |||  | |||| | ||
Db    662 CTGAATCGATCGAAATTATTGTAGCGCATGAAGTTTGAGTGAGGCTATGGTTTCCCA 721
Qy    661 ACGGATAAAAAAGTGTCAAACATCCAAGACATTTTGCCACTACTTGAGGAAGTTCTTAAA 720
            | ||||| || ||||| | |||||  |||||| |||||   | ||||| |||| ||  |
Db    782 AGTGAAAAGAAAATTCTAGTGATCCAATGTGACCCTGCCTTGAAAATTGCCCATGCT 841
Qy    721 ACCAACCGTCCATTACTCCATTATTGCCAGATGATGGTGAAACACACTTCAAACCCTT 780
            |  |||| |||  ||  |  || |  ||| | |||| |||  ||| || | |||||
Db    782 AGTGAAAAGAAAATTCTAGTGATCCAATGTGACCCTGCCTTGAAAATTGCCCATGCT 841
Qy    781 GTCTTGAAACAAGATTCGTTGGTACTTTCAAATGTGGTTGCTGTCTGTCAAAACGCGCCAGGATTTGGT 840
            || || || || |||| ||| ||| |||||| |||| || | | |||| ||||  || ||| |||| ||
Db    902 GTCTTGAATAGGCCTAAAGGTGGTGTCTGGAGTGTGGGGCCAAGGCTCAGGGTGTTGGT 961
Qy    841 GATCGTCGTAAACGTATGCTTGAAGACAATTGCTATCTTGACAGGTGGTACAGTGAGATTACA 900
            ||| | || ||  | ||| ||||| | || | || | || |  | |||  || | |||| ||
Db    902 GTCTTGAATAGGCCTAAAGGTGGTGTCTGGAGTGTGGGGCCAAGGCTCAGGGTGTTGGT 961
Qy    901 GAGGATCTAGGACT 914
            ||||| |  ||| |
Db   1022 GAAAAAAAGGATT 1035
Qy    902 GACAATAGAAAGAAACCAGCCTAAAGATAATGGCTATTGCTACTGGTGGGGGCCAAGGTTTTG 1021
```

COMMENT   Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGG
            High quality sequence stop: 716.

FEATURES   Location/Qualifiers
     source   1..878
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="AKA"
                     /db_xref="taxon:29729"
                     /clone="GA__Eb0024G23f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT   253 a   157 c   219 g   245 t   4 others
ORIGIN

Query Match   16.0%;   Score 261.8; DB 12;   Length 878;
Best Local Similarity 58.1%;   Pred. No. 6.6e-62;
Matches 458; Conservative   0; Mismatches 330; Indels   0; Gaps   0;

Qy   467 TTGGAGAGTATATCTCAGAAGCTATGGAGGCGTGTGGGCAACGATGGTGTGGGAACAG 526
          |||||| ||| |||||||| | ||||| ||| || |||| | || |||| ||| |||
Db   28 TTGGAACAATGATTGCTGATCGATTGCAAATGGTTGGACCTGTGGTGTTTGTCCCATTG 87

Qy   527 AAGAATCTCAGGTATAGGAAAACAGAACTTGAAGTGGTTGAAGGCATCCAAATTGACCGTG 586
          ||| | ||| || | | || ||| |||| | ||| || || || ||| ||| |||||
Db   88 AGTCATCATCCTCATTGAGACCACAGTTGATGTGAGGAAGGAAATGGAGATTGACACAG 147

Qy   587 GTTACCTGTCTCAATACATCATGGTCACAGACAATGAAAAAATGTTTGCAGACCTTGAAAAACC 646
          ||||| ||| ||| | || | | |||| ||| |||| | || ||| ||| |||| |||
Db   148 GTTACATCTCCCCACAAATTTGTTAATAAAACCCGGAGAAATTGGTTGTTGAGTTTGAGAATG 207

Qy   647 CATTTATCTTAATCACGGATAAAAAAGTGTTCAAAGACATCCAAGACACATTTTGCCACTACTTG 706
          |||| || || |||| ||| | ||||| ||| ||| |||| | ||| || ||||| ||
Db   208 CAAGACGTGTTGGTAACTGATCAAAAAGATTTCAGCTATAAAAGACATCATCCCCTGTTAG 267

Qy   707 AGGAGTTCTTAAAACAAACCGTCCATTACTCATTATTCCAGATAGTGATGGTGAAG 766
          ||||| || ||| ||| | || || || | ||| ||| ||| || ||||| |||
Db   268 AAAAGACCACTCAATTAAGATCTCCTTGCTTAATAATTGCTGAGGATGTGAGGAGAGG 327

Qy   767 CACTTCCAACCCTTGTCTTGAACAAGATTCGTGGTACTTTCAATGGTGGTTGCCTCTCAAAG 826
          |||||| || || || | ||| |||| | | ||| ||| ||| || || ||| ||||
Db   328 CTCTGGCCACACTTGTCGTGTGTGAACAAACTGCGTGGCATTCTCAATGTTGCAGCCATTAAG 387

Qy   827 CGCCAGGATTTGGTGATCGTCGTAAAGCTATGCCTTGAAGACATTGCTTATCTTGACAGGTG 886
          |||| ||| |||||||| ||| | | |||| ||| ||| |||| | || ||| ||||||||
Db   388 CTCCTGGTTTTGGTGAAAGGAGAAAAGCTCTCCTTCAAGATATTGCCATTCTGACTGGTG 447

Qy   887 GTACACTGATTACAGAGGACTTAGAGACTTGAATTAAAAGATGCTACACAAGAGCAGCCCCTTG 946
          | |||| ||| || |||| | | | | ||| || || | || ||| || | |||| ||
Db   448 CTGAGTTCCAAGCTAGTGCATTTGGGTTTCGTCGTCCAGAATACCTCAGTTGAGCAGCTTG 507

Qy   947 GACAGGCTGCTAGAGATTACACAGTTGATAAAGATAAGCACACGTAATTGTTGAAGGTTCAGGAA 1006
          || || | | ||| |||| || | || | ||| | | || ||| |||| ||||||||
Db   508 GTATTGCCAGAAGGTGATCATTACCAAGGAGATCGAACTCAACTAATTGCTGANGCAGCCT 567

Qy   1007 GTTCAGAAGGCTATTGCTAACCGTATTGCACTGATTAAAATCGCAAATTAGAAACAACAACTT 1066
          ||| ||| ||| | | |||||| ||||| |||| |||| ||||| ||| |||||| | ||
Db   568 CAAAGGATGAGAGTAACACTAGGGTGGCCAACTAAAAAAAAAGAGCTAGCTCAGACAGATT 627

Qy   1067 CTGACTTTGACCGTGAAAAAACTACAAAGAAACGTTGGCGAAATTGAGCAGCTG 1126
          |||| ||| || ||||||| | |||| ||| |||| |||| ||||||||
Db   628 CTGTCCTNATGATTCAGAAAAAATTGGCAGAAAAGGATTGCCAAACTATCTGTGGTGGTTCAG 687

Qy   1127 TTATCAAAGTAGGAGGAGCTCCAACAGACAGAGCAGCTTTAAAAGAAAATGAAAACTTCGGCATTGAGG 1186
          ||||| | || |||| | ||| | ||| || ||| || ||||||| || ||||| ||| | ||
Db   628

Qy   1127

Db   688 ATGCTCTATAAATGCTACACGTGCGCAGCCGGTTGAAGAAAGGTAGTCGTTGCGTGTGGTTGGGAACAG 747
           ||| || ||| | | || ||| || | ||| ||| | || || ||| || ||| ||||

Qy   1187 ATGCTCTAAATGCTACACGTGCGCAGCCGGTTGAAGAAAGGATGATCGTTGCGTGTGGTTGGAACAG 1246
           ||||| | |||||| | || ||| || | | || | || ||| || || |||||||

Db   748 ATGCAAAAGAATGCTACACAATTGCTGCCATAGNAGNAGAGGTATTGTGCCTGGGTGGGTGGTGCCTG 807

Qy   1247 CACTTATT 1254
          || | ||

Db   808 CCTTAGTT 815

RESULT 11

LOCUS   AL518632   1035 bp   mRNA   linear
DEFINITION AL518632 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA009Yl15 5'
            prime, mRNA sequence.
ACCESSION   AL518632
VERSION   AL518632.1 GI:12782125
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1035)
  AUTHORS   L1,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr.

FEATURES   Location/Qualifiers
     source   1..1035
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DA009Yl15"
                     /clone_lib="LTI_NFL011_NBC1"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library is not normalized, but is the control for
                     the normalized libraries. Library was constructed by Life
                     Technologies. Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"

BASE COUNT   311 a   195 c   259 g   266 t   4 others
ORIGIN

Query Match   15.7%;   Score 257.2; DB 9;   Length 1035;
Best Local Similarity 56.3%;   Pred. No. 1.4e-60;
Matches 500; Conservative   1; Mismatches 384; Indels   3; Gaps   1;

Qy   4 GCAAAAGAAATCAAATTTTCAGCAGAGATGCCGCGTTGCCGCCATGGGTGCCGCGGAGTTGATATG 63
          ||||||| | |||| || || || ||| || || | || || || ||| |||||| | |
Db   116 GCCAAAGANTGTAAAATTTGGGTGCAGATGCCCGAGCCTTAATGCTTCAAGGTGTAGACCTT 175

Qy   64 TTAGCAGAATACCGTCAAAGTAACCCTTGGTTCCTAAAAGGCGCCGCAATGTGTTCTTGAAAAA 123
          || || ||| ||| |||| || | || | | |||| || || || | || | ||| ||
Db   176 TTAGCCGATGCTGTCCCGTTACAATGGGCCCAAAGGGAAGAAGCCAAGGGAAGAACAGTGATTATTATTGAGCAG 235

Qy   124 GCTTTGGTTCTCCCTTAATTACTAATGAGGGGTAACCCATTGTTAAAGAGCATCGAATTA 183
          || || | |||| ||| | || ||| | || | || ||| || | |||| |||
Db   236 AGTTGGGGAAGTCCCAAAGTAACAAAAGAGTGGGTCGTGACTGTTGCAAAGTCAATTCAANTA 295

Qy   184 GAAGATCANTTTGAAAAACATGGGAGCAAAAATTGGGTGTCTGAAGTGGCTTCTAAAACCAAT 243
          ||| | ||| || | | || | | || | || | | | |||| || || ||||| |

```
Db  296 AAAGATAAAATACAAAAACATTGGAGCTAAACTTGTTCAATAACACAAAT 355
Qy  244 GATATTGCTGCTTGGGATGGGACGACTACTCGAACAGTTTGACACAAGCCATTGTTCATGAA 303
Qy  356 GAAGAAGCTGGGGGATGGCACTACCACCTGCTACTGTACTGGCACGCTCTATACCCAAGGAA 415
Db  304 GGACTAAAAAATGTGACACGCAGGGTGCTAATCCAATTGGTATCCGTACGGCCATTGAAACA 363
Qy  416 GCAACAGCAAACAGCTGTTGAAGGCCTTGAAAGCCCTTCAACCTGTATCTGGCAAGGAA 423
Db  356 GGCTTCGAGAAGATTAGCAAAAGGTGCTTAATCCACTGGAAAATCAGGACGAGGTGTGATGTTA 475
Qy  364 GCAACAGCAAACAGCTGTTGAAGGCCTTGAAAGCCCTTCAACCTGTATCTGGCAAGGAA 423
Db  416 GGCTTCGAGAAGATTAGCAAAAGGTGCTTAATCCACTGGAAAATCAGGACGAGGTGTGATGTTA 475
Qy  476 GCTGTTGATGTCGTGTAATTGCTGAACCTAAAAAGCAGTCTGAACCACCCCCTGAA 535
Db  476 GCTGTTGATGTCGTGTAATTGCTGAACCTAAAAAGCAGTCTGAACCACCCCCTGAA 535
Qy  424 GCTATTGCTCAGGTCCGCTCGAGTATCATCACGCTCTGAAAAAG---TTGGAGTGATATC 480
Db  536 GAAAATTGCACAGGTTGCTACGATTTCTGCAAACGACAAAGAAATTGGCAATATCATC 595
Qy  481 TCAGAAGCTATGGACGGTGCTGGGCAACGATGGTGGTGCAACGATGGTTGATTACCATCGA 540
Db  596 TCTGATGCAATGAAAAAGGTGTCATCACAGTAAAGGATGGGAAAAACA 655
Qy  541 ATGGAAACAGAACTTGAGAGTGGTTGAAGGCCATGCAATTTGACCGTGGTTACCGTCTCAA 600
Db  656 CTGAATGATGAATTAGAAATTATTGAAGGCCATGAAAGTTTGAATTGGCAATATCATCTCCA 714
Qy  601 TACATGGTCACAGACAATGAAAAAATGGTTGCAGAACCTTGAAAACCCATTTATCTTCAATC 660
Db  716 TACTTTATTAATACATCAAAAGGTCAGAAATGTCAGAGGCTATATTTCTG 775
Qy  661 ACGGATAAAAAAGTGTCAAACATCCAAGACACATTTCCCACTACTTACTTGGAAGTTCATAAA 720
Db  776 AGTGAAAAGAAAATTTCTAGTATCCATGGTGGCATGAAAGGTTGATTTCGCAAATGCCAAATGCT 834
Qy  721 ACCAAACCGTCCATTACTCATTATTTCCAGATGATGTGGAAGGTTCAAACCCTT 780
Db  836 CACCGTAAGCCTTTGGTCATAATCGCCTGAAGATGTTGATGGAGAAGCTTAAGTACACTC 894
Qy  781 GTCTTGAACAAGAATTCGTTGGTGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGAATTTGGT 840
```



```
BASE COUNT     306 a    188 c    242 g    251 t    1 others
ORIGIN

Query Match    15.5%;   Score 253.4;   DB 9;   Length 988;
Best Local Similarity  57.1%;   Pred. No. 1.6e-59;
Matches 500;   Conservative  0;   Mismatches 371;   Indels 4;   Gaps 2;

Qy    4 GCAAAAGAAAATCAAATTTTCACGACAGATGCGCGGTGCCATGGTGCGCGGAGTGATAG 63
Db  115 GCCAAAGATGCTAAAATTGGTGCAGATGCCGAGCCTTAATGCTTCAAGGTGTAGACCTT 174
Qy   64 TTAGCAGATACCGTCAAAGTAACGTGTTGGTCTCCCTAAAAGGGCGCCAATGTTGTTCTTGAAAA 123
Db  175 TTAGCCGATGCTGTGGCGGTTACAATGGGCCAAAGGCCAAAGGGAGAACAGTCAATTA 234
Qy  124 GCTTTTGGTTCTCCCTTAATTACTAATGACGGGGGTAACCATTGCTAAGAATGACGATCGAAATTA 183
Db  235 AGTTGGGGAAGTCCCAAAGTAACAAAAGATGGGGTGAGCTGTTGCAAAGTCAATGACTTA 294
Qy  184 GAAGATCATTTTGAAAACATGGGAGGCAAAAATTGGTGTGTCTGAAGTGGCTTCTAAAACCAAT 243
Db  295 AAAGATAAATACAAAAACATTGGAGCTAAACTTGTTGCCAATAACACAAAT 354
Qy  244 GATATTGCTGCTTGGGATGGGACGACTACTCGAACAGTTTGACACAAGCCATTGTTCATGAA 303
Db  355 GAAGAAGCTGGGGGATGGCACTACCACCTGCTAATTGGTTACTGTACTGGCAATGACTTA 414
Qy  304 GGACTAAAAAATGTGACACGCAGGGTGCTAATCCAATTGGTATCCGTACGGCCATTGAAACA 363
Db  415 GGCTTCGAGAAGATTAGCAAAAGGTGCTTAATCCACTGGAAAATCAGGACGAGGTGTGATGTTA 474
Qy  364 GCAACAGCAAACAGCTGTTGAAGGCCTTGAAAGCCCTTCAACCTGTATCTGGCAAGGAA 423
Db  475 GGCTTCGAGAAGATTAGCAAAAGGTGCTTAATCCACTGGAAAATCAGGACGAGGTGTGATGTTA 534
Qy  424 GCTATTGCTCAGGTCCGCTCGAGTATCATCACGCTCTGAAAAAG---TTGGAGTGATATC 480
Db  535 GAAAATTGCACAGGTTGCTACGATTTCTGCAAACGACAAAGAAATTGGCAATATCATC 594
Qy  481 TCAGAAGCTATGGACGGTGCTGGGCAACGATGGTGGTGCAACGATGGTTGATTACCATCGA 540
Db  595 TCTGATGCAATGAAAAAGGTGTCATCACAGTAAAGGATGGGAAAAACA 654
Qy  541 ATGGAAACAGAACTTGAGAGTGGTTGAAGGCCATGCAATTTGACCGTGGTTACCGTCTCAA 600
Db  655 CTGAATGATGAATTAGAAATTATTGAAGGCCATGAAAGTTTGAATTGGCAATATCATCTCCA 714
Qy  601 TACATGGTCACAGACAATGAAAAAATGGTTGCAGAACCTTGAAAACCCATTTATCTTCAATC 660
Db  715 TACTTTATTAATACATCAAAAGGTCAGAAATGTCAGAGGCTATATTTCTGTG 774
Qy  661 ACGGATAAAAAAGTGTCAAACATCCAAGACACATTTCCCACTACTTACTTGGAAGTTCATAAA 720
Db  775 AGTGAAAAGAAAATTTCTAGTATCCATGGTGGCATGAAAGGTTGATTTCGCAAATGCCAAATGCT 834
Qy  721 ACCAAACCGTCCATTACTCATTATTTCCAGATGATGTGGAAGGTTCAAACCCTT 780
Db  835 CACCGTAAGCCTTTGGTCATAATCGCCTGAAGATGTTGATGGAGAAGCTTAAGTACACTC 894
Qy  781 GTCTTGAACAAGAATTCGTTGGTGTACTTTCAATGTGGTTGCTGTCAAAGCGCCAGGAATTTGGT 840
```

RESULT 13

```
Db   895 GTCCTTGAATAGGCTAAAGGTTGGTCTTCAGGTTGTGGCAGTCAAGGCTTCAGGGTTGGT 954
         ||||||| || | | | |    | | | || |  |||||| || ||||| |||| |||||
Qy   841 GATGCTTCGTAAAGCTATTGCTTGAAGACATTGCCTAT 875
```

```
Db   955 GACAATAG-AAAGAMCAGCTTAAAGATATGGCTAT 988
```

BG321293
LOCUS   BG321293   977 bp   mRNA   linear   EST 27-FEB-2001
DEFINITION Zm04_04g09_R Zm04_AAFC_ECORC_cold_stressed_maize_seedlings Zea mays
            cDNA clone Zm04_04g09, mRNA sequence.
ACCESSION   BG321293
VERSION   BG321293.1 GI:13150971
KEYWORDS   EST.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 977)
  AUTHORS   Singh,J.A, Nakul,K., Couroux,P., De Moors,A., Harris,L.J., Hattori
            J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.
  TITLE   Expressed Sequence Tags from Cold-Stressed Maize Seedlings
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
            0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@em.agr.ca.
FEATURES   Location/Qualifiers
     source   1..977
                     /organism="Zea mays"
                     /cultivar="CO328"
                     /db_xref="taxon:4577"
                     /clone="Zm04_04g09"
                     /clone_lib="Zm04_AAFC_ECORC_cold_stressed_maize_seedlings"
                     /tissue_type="leaf, crown"
                     /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
                     Site_2: Xho I; Lower temperature 50 C / hour from 22 to
                     12oC; bring to 5o in 1 hour from 12oC. Leave at 5oC 2 days
                     ; photoperiod 16 hours. Light intensity was 125 uE-1.
                     library prepared by in vivo mass excision from amplified
                     library."
BASE COUNT   276 a   205 c   241 g   223 t   32 others
ORIGIN

Query Match   15.4%; Score 253; DB 12; Length 977;
Best Local Similarity   57.0%; Pred. No. 2.1e-59;
Matches 436; Conservative   22; Mismatches 304; Indels   3; Gaps   1;

```

```
Qy   4 GCAAAAGAAAATCAAAATTTTCAGACAATGCGCGTGCTGCCCGGAGTTGATATG 63
         || |||| || || ||| |||| ||| | ||| | |||||| | ||||| ||
Db   213 GCGAAGGACATCAAGTTTGGGTGTTGAGCGCCCGTTCTCYSAYGTTGAAGGTGTTGAGGAG 272
```

```
Qy    64 TTAGCAGATACCGTCAAAGTAACGCTTGGTCCTCTAAAAGGGCGCAAATGTTGTTCTTGAAAAA 123
         || | |||  |||| |  || ||||| | |||  | || ||||| |||||||||||||| ||||
Db   273 TTGGCAGAWGCTGTTAAAGTTACAATTGGACCTAAGGGCGCCGAAATGTKGTTATTGAGGCAA 332
```

```
Qy   124 GCTTTTGGTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGAGATCGAATTA 183
         |||:||| ||      |  | ||||| | ||| || | ||| | | |||||||||| ||:
Db   333 AGCTTTTKGTCGACCGAAAGTCACAAAGGATGGTGTTYACTTWAGCAAAGAGACCATTGAATTY 392
```

```
Qy   184 GAAGATCATTTTGAAAACATGGGAGCAAAAAATTGGTGTCTGAAGTGGCCTTCTAAAACCAAT 243
         ||| || ||  |||| |||  ||| || |||| || || ||   || | |||| ||:|||
Db   393 AGGATAGAGTAAAAGAANTGTTGGTGTGCCAAGCSTTGTGAAACAGGTTGCTTAATGCWACYAYAT 452
```

```
Qy   244 GATAATTGCTGGGTCGGTGATGGGACGACTACTCCCAACAGTTTTGACACAAGCCATTGTTCATGAA 303
         || | |||||| ||| ||| | | ||| || ||  | ||  ||  |:|:||| |  || |||
Db   513 ...
```

RESULT 14

```
Db   453 GACACTTCYGGGGATGGTACCACATGTKCCACTKTTTTGACAAAGGACAAATATTTACTCGAG 512
         ||     | ||   |||||  ||||||  |||| ||||| |||| | ||||||||||||| ||
Qy   841 ...
```

AL515262
LOCUS   AL515262   955 bp   mRNA   linear   EST 13-FEB-2001
DEFINITION AL515262 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0B0162A04 5
            prime, mRNA sequence.
ACCESSION   AL515262
VERSION   AL515262.1 GI:12778755
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 955)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE   Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES   Location/Qualifiers
     source   1..955
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CL0B0162A04"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT   288 a   181 c   236 g   247 t   3 others
```

```
Db   304 GGACTAAAAAAYTGGACAGCACAGGTGCTTAATCCAATWTGTTAATCCGTTCGGTGCAGGCGATTGAAAC 363
         ||   |||||||| |||||| | |||| || ||| ||||| || ||||||| ||   | |||||||
Qy   ...
```

```
Db   513 GGGTGCAAATCTGTTGCGCGGGCTGGAATKAATGCTATGGATTTWAAGGCGTGGAAACTCAATG 572
```

```
Db   364 GCAACAGAAACAGCTCGTTCAAGGCCTTGAAAGCCATTGCTCAACCTGTATCTGGCAAGGAA 423
         |  |  | | | |  || |||| | |||||| ||||||| |||   | ||||| |||| |||
Qy   573 GCTGTTGACGCGTGTTGTTACCAATCTTAAAARGCATGGCCAGAATKATCAGCACTTCAGAA 632
```

```
Qy   753 CTTTTATAATGAGCTTTAAGCTTGGTATTAAAACTAGACAGARGTTACACATCTCCG 812
```

```
Qy   541 ATGGAAACAGAACTTGAAGTGGTTGAAGGCCATGCAAFTTGACCGGTTACCTGTTCTCAA 600
```

```
Qy   481 TCAGAAGCTATGGAGCGGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCGAGGT 540
```

```
Qy   693 GCCAAGGCTATGGAGAAAGGTTGGCAAAGAGGGTGTYAWCACTATTGCGGACGGYAAACACC 752
```

```
Qy   424 GCTATTGCTCAGGTCGGTCGCCCAGTATCATCACCGTCTGAAA--AAGTTGGAGAGTATATC 480
```

```
Qy   633 GAAATTGCACAGGTCGGGTYACAAATATCAGCAAATATGGGGAAAAGGGAAAATTGGTAGGCBRATT 692
```

```
Qy   601 TACAAGGTCACAGACAAYTGAAAAAAAATGGTTGCAGAGACCTTGAAAAACCCATTTAYCTTRATC 660
```

```
Qy   813 TACTTCATTACCAACCCAAAAAAACCCAGAAAAYGTGAATTGGAAGACBCABYDCAYTCTTRATT 872
```

```
Qy   661 ACGGATAAAAAAGGTTCAAAACATCCAAGACACATTTGCCCACTACTTGAGGAAGTTCTTAAA 720
```

```
Qy   873 CATGACAAGAAGGTGACCAACATGCAATYCYGYGKTYYGKGTTAAGGKTTAGAAATGGCTCTYAAG 932
```

```
Qy   721 ACCAACCGTCCATTACTCATTATTGCCAGATGATGGACGTGGAATGGGATGGTGAA 765
```

```
Qy   933 AAACAAAAGCCTCWCTWCTGATTTTTTGCCAGAAGAAGTTGAAAGTTAA 977
```

```
RESULT 15
LOCUS         AL515579
DEFINITION    AL515579 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA002YD20 5
              prime, mRNA sequence.
ACCESSION     AL515579
```

```
LOCUS         AL515579                1017 bp    mRNA        linear    EST 13-FEB-2001
DEFINITION    AL515579 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA002YD20 5
              prime, mRNA sequence.
ACCESSION     AL515579
VERSION       AL515579.1  GI:12779072
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 1017)
AUTHORS       Li,W.B., Gruber,C., Jessee,J. and Polyes,D.
TITLE         Full-length cDNA libraries and normalization
JOURNAL       Unpublished (2001)
COMMENT       Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..1017
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DA002YD20"
                     /clone_lib="LTI_NFL011_NBC1"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer, five prime end
                     enriched, double-stranded cDNA was digested with NotI and
                     cloned into the NotI and EcoRV sites of the pCMVSPORT 6
                     vector. Library is not normalized, but is the control for
                     the normalized libraries. Library was constructed by Life
                     Technologies. Contact: Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      308 a    191 c    255 g    261 t      2 others
ORIGIN
```

```
Query Match          15.2%;   Score 249.6;   DB 9;  Length 1017;
Best Local Similarity  56.8%;   Pred. No. 1.9e-58;
Matches    478;  Conservative  1;  Mismatches 360;  Indels  3;  Gaps  1;
```

```
Qy     4 GCAAAAGAAATCAAAATTTTCAGCAGATACGCGTCGCCATGCGTCGCGGGAGTTGATATG 63
         || |||||| ||||||||||||||| ||| ||||||||| | | || |||||||||| |||
Db   116 GCCAAAGATGTAAAATTTGGTCGCAGATGCCCGAGCCTTAATCGCTTCAAAGGTGTAGACCTT 175

Qy    64 TTAGCAGATACCGTCAAAGTAACCGCTTGGTCCTTAAAGGCGGCCAATGTTGTTCTTGAAAA 123
         |||||||||||||||||||||| ||| |||| ||||||||||||| ||||||| ||||||
Db   176 TTAGCCGATGCTCTGGCGTTACAAATGGGGCCAAAAGGGAGAAGACAGTGATTATTGAGCAG 235

Qy   124 GCTTTTGCTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGATCGAATTA 183
         ||||| |||| || || ||||| || ||| |||| |||| |||||| || |||||
Db   236 AGTTGGGGAAGTCCCAAAGTAACAAAAAGATGGTGTGACTKTTGCAAAGTCAATTGACTTA 295

Qy   184 GAAGATCATTTTGAAAACATGGGAGCAAAATTGGGTGTCTCGAAGTGGCTTCTAAAACCAAT 243
         |||||| || |||| || ||||||||| ||||||| ||||| |||||||||| ||||||||
Db   296 AAAGATAAATACAAAAACATTGGAGCTAAAACTTGTTCAAGATGTTGCCAATAACACAAAT 355

Qy   244 GATATTGCTGGTGATGGAGACGACTACTGCAAACAGTTTTGACACAAGCCAATTGTTCATGAA 303
         ||||||||||||||||| |||||||| |||| || |||||||| ||||||||||||| ||
Db   356 GAAGAAGCTGGGGAATGGCACTACCACTGCTTACTGTACTGGCCACGCTCTATAGCCAAGGAA 415

Qy   304 GGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATNCCGTCGAAGGCATTGAAACA 363
         |||||||||||||||||| || ||| || |||||||| ||| || ||||||||| |||||
Db   416 GGCTTCGAGAAGATTAGCAAAGGTGCTAATCCAGTGGAAATCAGGAGAGGTGTGAGTTA 475

Qy   364 GCAACACAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCCTGTATCTGGCAAGGAA 423
         ||| | || |||| ||||||||||| |||| |||||||| || |||| |||||||||||
Db   476 GCTATTGCTCAGGTCGCCTGCAGGTATCATCACGCTCTGAAAAAG---TTGGAGAGTATATC 535

Qy   424 GCTATTGCTCAGGTCGCCTGCAGTATCATCACGCTCTGAAAAAG---TTGGAGAGTATATC 480
         ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   536 GAAATTGCACAGGTTGCTACGGATTTCTGCAAACGGAGACAAAGAAAATTGGCAATATCATC 595
```

```
Query Match          15.4%;   Score 251.6;   DB 9;  Length 955;
Best Local Similarity  57.0%;   Pred. No. 5e-59;
Matches    480;  Conservative  0;  Mismatches 359;  Indels  3;  Gaps  1;
```

```
Qy     4 GCAAAAGAAATCAAAATTTTCAGCAGATACGCGTCGCCATGCGTCGCGGGAGTTGATATG 63
         || |||||| ||||||||||||||| ||| |||||| ||| | | || |||||||| |||
Db   114 GCCAAAGATGTAAAATTTGGTCGCAGATGCCCGAGCCTTAATCGCTTCAAAGGTGTAGACCTT 173

Qy    64 TTAGCAGATACCGTCAAAGTAACCGCTTGGTCCTTAAAGGCGGCCAATGTTGTTCTTGAAAA 123
         ||||||||| |||||||| || ||| |||| ||||||||||||| ||||||| |||||||
Db   174 TTAGCCGATGCTCTGGCGTTACAAATGGGGCCAAAAGGGAGAAGACAGTGATTATTGAGCAG 233

Qy   124 GCTTTTGCTTCTCCCTTAATTACTAATGACGGGGTAACCATTGCTAAAGATCGAATTA 183
         ||||| |||| || || ||||| || ||| |||| |||| |||||| || |||||
Db   234 AGTTGGGGAAGTCCCAAAGTAACAAAAAGATGGTGTGACTKTTGCAAAGTCAATTGACTTA 293

Qy   184 GAAGATCATTTTGAAAACATGGGAGCAAAATTGGGTGTCTCGAAGTGGCTTCTAAAACCAAT 243
         |||||| || |||| || ||||||||| ||||||| |||||| |||||||||| ||||||||
Db   294 AAAGATAAATACAAAAACATTGGAGCTAAAACTTGTTCAAGATGTTGCCAATAACACAAAT 353

Qy   244 GATATTGCTGGTGATGGAGACGACTACTGCAAACAGTTTTGACACAAGCCAATTGTTCATGAA 303
         ||||||||||||||||| |||||| ||||| || ||||||| | ||||||||||||| ||
Db   354 GAAGAAGCTGGGGATGGCCACTACCACTGCTTACTGTACTGGCCACGCTCTATAGCCAAGGAA 413

Qy   304 GGACTAAAAAATGTGACAGCAGGTGCTAATCCAATTGGTATNCCGTCGAAGGCATTGAAACA 363
         |||||||||||||||||| ||| || || |||||||| ||| || ||||||||| |||||
Db   304 GGCTTCGAGAAGATTAGCAAAGGTGCTAATCCAGTGGAAATCAGGAGAGGTGTGAGTTA 363

Qy   364 GCAACACAACAGCTGTTGAAGCCTTGAAAGCCATTGCTCAACCCTGTATCTGGCAAGGAA 423
         ||| | || |||| ||||||||||| |||| |||||||| || |||| |||||||||||
Db   414 GGCTTCGAGAAGATTAGCAAAGGTGCTAATCCAGTGGAAATCAGGAGAGGTGTGAGTTA 473

Qy   424 GCTATTGCTCAGGTCGCCTGCAGTATCATCACGCTCTGAAAAAG---TTGGAGAGTATATC 480
         ||||||||||||| | ||||||||||||||||||| ||||| | |||||||||||||||
Db   474 GCTGTTGATGCTGTTTATTGCTGAACTTAAAAAGCAGTTTAAACCTGTTAAACCCCCTGAA 533

Qy   481 TCAGAAGCTATGGAGCGGTGTGGGCAACGATGGTGTGTGATTACCATCGAAGAATCTCGAGGT 540
         ||||| || ||||||| ||||| ||||| |||||||| || ||| |||||||| ||||||
Db   424 GAAATTGCACAGGTTGCTACGGATTTCTGCAAACGGAGACAAAGAAAATTGGCAATATCATC 480

Qy   541 ATGGAAACAGCAACTGTGAAGCTGGTTGAAGGCATGCAATTGACCGTGGTTACCGTGGTTGATATG 600
         ||||||| | |||| |||| || ||||||| ||| ||||| |||| || |||| ||||||
Db   594 TCTGATGCAATGAAAAAGTTGGAAGAAAAGGTGTTCATCACAGTTAAGGATGGAAAAACA 653

Qy   601 TACATCGTCACGACAATGAAAAAATGGTTGCAGACCTTGAAAACCCATTATCTTAATC 660
         ||| | || | || ||||| || |||||| |||||||| ||||||||| |||||||
Db   654 CTGAATGATGGAATTATTGAGATGAGTTCATGCAATTAACTCTTAACGTTATGTAGCTTG 713

Qy   661 ACGGATAAAAAACTGTCAAACACTCCAAGACATTTGCCACTACTTGAGGAAGTTCTTAAA 720
         ||||||||||||| |||||| ||| ||||| |||| |||| |||| ||||||| ||||
Db   714 TACTTTATTAATACATCAAAAGGTCAGAAAATGTGAATTCCAGGATGCCTATNGTTCGTCGTG 773

Qy   721 ACCAACCGTCCATTACTCATTATTGCAGATGCAGTGGATGATGGATTGAAGCACTTCCAACCCTT 780
         ||| || || ||||||| ||| || ||||| || ||||||| ||||| ||| ||| ||||
Db   774 AGTGAAAAAGAAAAATTCTAGTATCCAATTGTACCTGCTCTTGAGGKTGCCCAATGCT 833

Qy   781 GTCTTCAACAAGATTCGTGTACTTCAATCTGGTTGCTGTCAAAGCGCCAGGAATTTGGT 840
         |||||||| |||| || |||| || || |||||||| ||||||||||| || |||||
Db   834 CACCGTAAGCCTTTGGTCATAATCGCTGAAGATGTTGATGGAGAAGCTCTAGCTACACTC 893

Qy   841 GTCTTGAATAGGCTAAAGGTTGGGTTCTTCAGGTTGTTGGGCAGTAAAGGCTCCAGGGTTTGGT 894
         ||| || |||| |||| || || ||||||||||||||||| |||||||||| || ||||||
Db   894 GTCTTGAATAGGCTAAAGGTTGGGTTCTTCAGGTTGTTGGGCAGTAAAGGCTCCAGGGTTTGGT 953

Qy   841 GA 842
         ||
Db   954 GA 955
```

```
Qy  481  TCAGAAGCTATGGACGGTGTGGGCAACGATGGTGTGATTACCATCGAAGAATCTCGAGGT 540
         || | || | | |-| | | |-| ||| || | || || | | | || | | || | |
Db  596  TCTGATGCAATGCAAAAAAGTTGGAGAAAAGGTGTCATCACAGTAAAGGATGGGAAAAACA 655

Qy  541  ATGGAAACAGAACTTGAAGTGGTTGAAGGCATGCCAATTTGACCGTGGTTACCTGTCTCAA 600
         ||| | ||| | || || | ||||||||| | | ||||| | || || | || || | |-
Db  656  CTGAATGATGAATTAGAAATTATTGAAGGCATGAAGTTTGATCGAGGCTATATTTCTCCA 715

Qy  601  TACATGGTCACAGACAATGAAAAAAATGGTTGCAGACCTTGAAAACCCATTTATCTTAATC 660
         |||-| | || |-| | | |-| |-| | | | |-| |-|
Db  716  TACTTTATTTAATTAATACATCAAAAAGGTCAGAAAATGTGAATTCCAGGATGCCTATGTTC 775

Qy  661  ACGGATAAAAAAGTGTCAAACATCCAAGACACATTTTGCCACTACTTGAGGAAGTTCTTAAA 720
         | -| | || | |-| | | | | |||||  |||| | -|
Db  776  AGTGAAAAAGAAAAATTTCTAGTATCCAATTCCAGTCCATTGTACCTGCTCTTGAAATTGCCAAATGCT 835

Qy  721  ACCAACCGTCCATTACTCCATTATTATTGCAGAATGATGATGGTGAAGCACTTCCAACCCTT 780
         | -| | | |-| | || | || | | | | | | | | || |  |
Db  836  CACCGTAAGCCTTTGGTCATAATCCGCTGAAGATGTCGATGGAGAAGCTCTAAGTACACTC 895

Qy  781  GTCTTGAACAAGATTCGTGGGTACTTTCAAATGTGGTTGCGTCTCAAAGCGCCAGGAATTGGT 840
         ||||||| | -| | | | | | | | | |-| ||||| | || |||||
Db  896  GTCTTGAATAGGCTAAAGGTTGGTCTTCAGGTTGTGGCAGTAAAAGGCTCCAGGGTTTGGT 955

Qy  841  GA 842
         ||
Db  956  GA 957
```